Model Reduction by Balanced Truncation

submitted by

Chris Guiver

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

January 2012

Signature of Author ....

Chris Guiver

Summary

Model reduction by balanced truncation for bounded real and positive real input-state-output systems, known as bounded real balanced truncation and positive real balanced truncation respectively, is addressed. Results for finite-dimensional systems were established in the mid to late 1980s and we consider two extensions of this work. Firstly, using a more behavioral framework we consider the notion of a finite-dimensional dissipative system, of which bounded real and positive real input-state-output systems are particular instances. Specifically, we work in a framework where we make no a priori distinction between inputs and outputs. We derive model reduction by dissipative balanced truncation, where a gap metric error bound is obtained, and demonstrate that the aforementioned bounded real and positive real balanced truncation can be seen as special cases.

In the second part we generalise bounded real and positive real balanced truncation to classes of bounded real and positive real systems respectively that have non-rational transfer functions, so called infinite-dimensional systems. Here we work in the context of well-posed linear systems. We derive approximate transfer functions, which we prove are rational and preserve the relevant dissipativity property. We also obtain error bounds for the difference of the original transfer function and its reduced order transfer function, in the H-infinity norm and gap metric for the bounded real and positive real cases respectively. This extension to bounded real and positive real balanced truncation requires new results for Lyapunov balanced truncation in the infinite dimensional case, which we also describe. We conclude by highlighting possible future research.

Acknowledgements

I would like to thank my supervisor Mark Opmeer for being a fantastic supervisor. This project would not have been possible without him and I am very grateful for all his support, advice, effort, time, energy and skill.

I would also like to thank the control group at the University of Bath, mentioning particularly Hartmut Logemann and Gene Ryan, for their helpful and interesting contributions. I am grateful to Alastair King and Fran Burstall for their advice and for helping me develop as a mathematician. Thanks also go to the administrative and support staff in the Department of Mathematical Sciences for all their help along the journey and to the EPSRC who funded this project.

I would also like to thank my friends, remembering especially the 1W crew of Adam, James, Phil, Jane and Caz. Thank you all very much for all the maths help and much fun. Special mentions also go my housemate Nick who has been a great friend and my family for their love, patience and support.

Most of all I would like to thank God, for being awesome, bringing me to where I am today and making this project possible. Thank you for Your love, patience and mercy :) God is amazing and it is to Him that this thesis is dedicated.

List of symbols

$\mathbb{N}, \mathbb{N}_0$	set of natural numbers and set of non-negative integers.
$\mathbb{Z}, \mathbb{R}, \mathbb{C}$	ring of integers, field of real numbers and field of complex numbers.
$\mathbb{C}_\sigma^+$	open right half-plane $\{s \in \mathbb{C} : \operatorname{Re} s > \sigma\}$.
$\mathbb{C}_\sigma^-$	open left half-plane $\{s \in \mathbb{C} : \operatorname{Re} s < \sigma\}$.
$\mathbb{D}$	complex unit disc $\{s \in \mathbb{C} : \lvert s \rvert < 1\}$.
$\mathbb{T}$	complex unit circle $\{s \in \mathbb{C} : \lvert s \rvert = 1\}$.
$\mathbb{R}^+$	set of non-negative real numbers $\{x \in \mathbb{R} : x \geq 0\}$.
$\mathbb{R}^-$	set of non-positive real numbers $\{x \in \mathbb{R} : x \leq 0\}$.

For $\mathscr{X}, \mathscr{Z}$ Banach spaces.

$B(\mathscr{X}, \mathscr{Z})$	set of bounded linear operators $\mathscr{X} \to \mathscr{Z}$.
$B(\mathscr{X})$	set of bounded linear operators $\mathscr{X} \to \mathscr{X}$.
$D(A)$	domain of the linear operator A.
$\mathcal{G}(A)$	graph of the linear operator A.
$\rho(A), \sigma(A), \sigma_p(A)$	resolvent set, spectrum and point spectrum of the linear operator A.
$\sigma_+(A), \sigma_-(A)$	for $A : \mathscr{X} \to \mathscr{X}$ self-adjoint and $\mathscr{X}$ finite-dimensional, the number of positive and negative eigenvalues of A, counting multiplicities.

For $1 \leq p \leq \infty$.

$\ell^p(\mathscr{X})$	space of p-summable sequences $\mathbb{N} \to \mathscr{X}$.
$L^p(I; \mathscr{X})$	Lebesgue spaces of measurable, p-integrable functions $\mathbb{R} \supseteq I \to \mathscr{X}$.
$L^p_{\mathrm{loc}}(I; \mathscr{X})$	space of locally L^p functions $\mathbb{R} \supseteq I \to \mathscr{X}$.
$H^p(D; \mathscr{X})$	Hardy spaces of functions $\mathbb{C} \supseteq D \to \mathscr{X}$.

Contents

Chapter 1

Introduction

Model reduction for control systems refers to approximating a given model by a simpler one that is close to the original in some sense. Model reduction is important for simulation and controller design. There are many model reduction schemes in the literature, and this thesis focuses on balanced truncation.

Approximation of a transfer function by truncation of a balanced state-space realisation was first suggested for rational functions by Moore in [52]. A (Lyapunov) balanced realisation is a realisation where the controllability and observability Gramians are equal. The importance of balanced truncation relies on an explicit H^∞ bound on the difference of the transfer functions established independently by Enns [23] and Glover [26],

$$\|G - G_n\|_{H^\infty} \leq 2 \sum_{k=n+1}^{N} \sigma_k. \tag{1.1}$$

In the above inequality σ_k are the singular values of the Hankel operator associated to G, n is the order of the balanced truncated system with transfer function G_n and N is the order of the original system.

Many mathematical models incorporate the loss or dissipation of energy in their design. For example many physical systems such as mass, spring, damper systems or resistor, inductor, capacitor circuits are dissipative with respect to a certain supply rate; typically the (quadratic) scattering supply rate or the impedance supply rate, using the terminology of Willems [99]. Systems which are dissipative with respect to the former are often called scattering passive, contractive or bounded real. Systems which are dissipative with respect to the latter are often called impedance passive, passive or positive real. Although Lyapunov balanced truncation retains stability of the system, any energy relation that the original system satisfies may not be retained by the truncated system. This led to the introduction of bounded real balanced truncation by Opdenacker & Jonckheere in [57], where also an H^∞ error bound is provided. Earlier positive real balanced truncation was introduced by Desai & Pal [22]. There exist error bounds in that case too [3], [34] but these seem somewhat less natural than the error bounds for the other balanced truncation methods (a case in point being that this is the only balanced truncation method for which there are several different error-bounds in the literature and a false bound in [14]). These two approximation methods preserve the relevant energy relation in the sense that the transfer function of the truncated system is again bounded real (respectively, positive real).

The aim of this thesis is to extend bounded real balanced truncation and positive real balanced truncation. We do so in two ways. In Part I we extend the concept of bounded real and positive real input-state-output systems to that of a dissipative system. That bounded real and positive real input-state-output systems are essentially the same system looked at in different ways is of course well-known: positive real systems and bounded real systems are related by a transform which goes by the different names of Cayley transform, Möbius transform and diagonal transform. However, this relationship follows in a more natural way by considering these systems in a behavioral [67] or state/signal [4], [44], [100] framework. In these frameworks no a priori distinction is made between inputs and outputs. Instead, an external signal is studied which may be decomposed into an input and an output in various ways.

It is argued by Willems [101] that the choice of inputs and outputs of a given system is often artificial, and so we apply model reduction by balanced truncation in the above framework which is free of such constrictions. We derive an error bound for the distance between a dissipative system and its dissipative balanced truncation in the gap metric. The gap metric is a behavioral object in the sense that it does not depend on the input/output decomposition. Bounded real systems and positive real systems appear when specific input/output decompositions are chosen in the same dissipative behavioral or state/signal system. By choosing specific input/output decompositions we recover bounded real and positive real balanced truncation and obtain a new error bound in the gap metric for positive real balanced truncation

In Part II, we extend bounded real and positive real balanced truncation in a novel way by generalising the theory to non-rational transfer functions. In this case any realisation must have an infinite-dimensional state-space. Using the Cayley transform we obtain results for positive real balanced truncation from their bounded real balanced truncation counterparts. We demonstrate that under certain assumptions the bounded real and positive real balanced truncations exist, preserve the respective dissipativity property and the corresponding finite-dimensional error bounds extend to their infinite-dimensional counterparts.

In the finite-dimensional case, bounded real balanced truncation can be seen as Lyapunov balanced truncation of a certain extended system. We mimic this approach in the infinite-dimensional case. Here the existence of Lyapunov balanced realisations is non-trivial. A special case was treated in Curtain & Glover [17] and the general discrete time case was proven by Young [103]. This was subsequently converted to general continuous-time systems by Ober & Montgomery-Smith [56]. In Glover *et al.* [27] balanced truncations and the H^∞ error bound (with $N = \infty$ in (1.1)) were extended to a class of infinite-dimensional systems.

In order to derive bounded real balanced truncation in the infinite-dimensional case we needed to extend the above existing Lyapunov balanced truncation results to a larger class of systems. These results are also of independent interest.

This thesis is organised as follows. Part I considers finite-dimensional theory. We begin in Chapter 2 with a review of model reduction by balanced truncation (in the finite-dimensional case), recalling Lyapunov, bounded real and positive real balanced truncation. In Chapter 3 we generalise these notions to that of dissipative systems and dissipative balanced truncation. We begin Part II by collecting some elementary infinite-dimensional systems theory and specific functional analysis results that we will need. Chapters 5, 6 and 7 consider Lyapunov, bounded real and positive real balanced truncation in the infinite-dimensional case respectively. In Chapter 8 we provide some easily verifiable sufficient conditions for when the main results from Chapters 6 and 7 are applicable and include an example. The final chapter contains some summarising remarks and suggestions for future work. Chapters 3 and 5-8 contain most of the new material in this thesis and these chapters begin with a short outline motivating and highlighting the results of that chapter. Every chapter concludes with a notes section that contains more relevant background material, mentions what is novel and what we have published or submitted for publication.

Part I

Finite-dimensional theory

Chapter 2

Review of finite-dimensional model reduction by balanced truncation

Let $\mathscr{U}$ and $\mathscr{Y}$ denote finite-dimensional Hilbert spaces, which are the input and output spaces respectively. We recall that a rational function $G : \mathbb{C}_0^+ \to B(\mathscr{U}, \mathscr{Y})$ belongs to H^∞ if and only if G is proper and every pole of G is in the open left-half complex plane. Given such a G then it is possible to write

$$G(s) = D + C(sI - A)^{-1}B, \quad s \in \mathbb{C}_0^+, \tag{2.1}$$

for some finite-dimensional space $\mathscr{X}$ and operators

$$A : \mathscr{X} \to \mathscr{X}, \quad B : \mathscr{U} \to \mathscr{X}, \quad C : \mathscr{X} \to \mathscr{Y}, \quad D : \mathscr{U} \to \mathscr{Y}, \tag{2.2}$$

with A stable (that is, every eigenvalue has negative real part, also known as Hurwitz). The quadruple of operators (2.2) (and implicitly the space $\mathscr{X}$) is called a realisation of G and is denoted by $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. Moreover, we always choose $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ such that the associated input-state-output system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \\ x(0) &= x_0, \end{aligned} \tag{2.3}$$

is minimal (i.e. controllable and observable). With the above construction G is the transfer function of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ or (2.3) in the usual way.

Model reduction typically seeks to approximate a transfer function G (or some other input-output object, such as the input-output map) with a simpler one. For example, by moment matching (interpolation) of the original transfer function. All model reduction by balanced truncation schemes acheive this aim by truncating states from a state-space realisation of G that are unimportant in some sense. Truncation in the state-space is dependent on the particular realisation which is chosen. Therefore it is crucial to be able to quantifying importance of states, in a manner that is independent of the realisation chosen.

2.1 Lyapunov balanced truncation

Lyapunov balanced truncation is a balanced truncation scheme for stable (that is, H^∞) transfer functions, based on the combined controllability and observability of states and as such makes use of the controllability and observability Gramians.

Definition 2.1.1. The controllability Gramian $\mathcal{Q}$ and observability Gramian $\mathcal{O}$ of a stable, minimal realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ are given by

$$\mathcal{Q} := \int_{\mathbb{R}+} \mathrm{e}^{At} B B^* \mathrm{e}^{A^* t} \, dt, \quad \mathcal{O} := \int_{\mathbb{R}+} \mathrm{e}^{A^* t} C^* C \mathrm{e}^{At} \, dt, \tag{2.4}$$

where e^A denotes the matrix exponential of A.

The operators $\mathcal{Q}$ and $\mathcal{O}$ are both self-adjoint, positive and thus invertible (since the realisation is minimal), but do depend on the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$.

It is well-known that the reachable subspace of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is equal to the image of $\mathcal{Q}$. Similarly, the unobservable subspace of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is equal to the kernel of $\mathcal{O}$. Although we have excluded this case by assuming our realisations are minimal, if $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ was *not* controllable, $\mathcal{Q}$ would have a zero eigenvalue. Similarly, if $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ was *not* observable then $\mathcal{O}$ would have a zero eigenvalue. One could argue that uncontrollable states are good candidates to omit in a reduced order system, as they are superfluous, and similarly for unobservable states. However, the notions of controllability and observability are independent in the sense that states which are unobservable may not be uncontrollable and vice versa. Even in the minimal case, heuristically, by continuity it does not seem unreasonable to expect that small eigenvalues of $\mathcal{Q}$ correspond to eigenspaces that are "nearly" uncontrollable (in some sense). These states again seem a good candidate to omit in a reduced order system (and similarly for small eigenvalues of $\mathcal{O}$). These observations motivate the following definition.

Definition 2.1.2. Let $\mathcal{Q}$ denote the controllability Gramian and $\mathcal{O}$ denote the observability Gramian of the minimal, stable realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. We say that $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is output-normal if

$$\mathcal{O} = I,$$

I the identity on $\mathscr{X}$. We say that $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is Lyapunov balanced, or in Lyapunov balanced co-ordinates, if

$$\mathcal{Q} = \mathcal{O} =: \Pi. \tag{2.5}$$

Balanced in the context of Lyapunov balanced truncation means obtaining a realisation where "states that are difficult to reach are simultaneously as difficult to observe." We expand on this shortly. A crucial result is the following.

Lemma 2.1.3. *Given a minimal, stable realisation* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *there exists an invertible transformation* T *such that* $\left[\begin{smallmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{smallmatrix}\right]$ *is Lyapunov balanced.*

Proof. See Antoulas [3, Lemma 7.3]. □

Remark 2.1.4. Computing Lyapunov balanced realisations involves firstly finding the controllability Gramian $\mathcal{Q}$ and observability Gramian $\mathcal{O}$ of a stable, minimal realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. As is well-known, $\mathcal{Q}$ and $\mathcal{O}$ are the unique solutions of the controller and observer Lyapunov equations

$$\begin{aligned} A\mathcal{Q} + \mathcal{Q}A^* + BB^* &= 0, \\ A^*\mathcal{O} + \mathcal{O}A + C^*C &= 0, \end{aligned}$$

respectively.

It turns out that the important quantities in Lyapunov balanced truncation are the Lyapunov singular values, which we describe now. The Hankel operator $H : L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$ of the transfer function (2.1) is given by

$$(Hf)(t) = \int_{\mathbb{R}^+} C\mathrm{e}^{A(t+s)} Bf(s)\, ds = (\mathfrak{C}\mathfrak{B}Rf)(t),$$

where R is the reflection in time $Rf(t) = f(-t)$ and (at least formally)

$$\mathfrak{B}u = \int_{\mathbb{R}^-} \mathrm{e}^{A(-s)} Bu(s)\, ds, \quad (\mathfrak{C}x)(t) = C\mathrm{e}^{At}x.$$

The maps $\mathfrak{B}$ and $\mathfrak{C}$ are the input and output map of the well-posed linear system generated by the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. We recall the notion of a well-posed linear system in Chapter 4 as we will not make use of them until Part II. The singular values of an operator are also defined precisely in Chapter 4. For now we assume that they are the non-negative square roots of the eigenvalues of H^*H (see Remark 4.3.3).

Lemma 2.1.5. *Let* H, $\mathcal{Q}$ *and* $\mathcal{O}$ *denote the Hankel operator, controllability and observability Gramians respectively of the transfer function with realisation* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. *Then* H^*H *and* $\mathcal{Q}\mathcal{O}$ *have the same non-zero eigenvalues.*

Proof. It is well-known that for $\mathscr{X}$ a Hilbert space and $S, T : \mathscr{X} \to \mathscr{X}$ bounded operators, $\lambda \neq 0$ satisfies

$$\lambda \in \sigma(ST) \iff \lambda \in \sigma(TS),$$

(contained in [60, Lemma 3.16] for example). Therefore for $\lambda \neq 0$, noting that

$$\mathcal{Q} = \mathfrak{B}\mathfrak{B}^* \quad \text{and} \quad \mathcal{O} = \mathfrak{C}^*\mathfrak{C},$$

we have that

$$\begin{aligned} \lambda \in \sigma(H^*H) &\iff \lambda \in \sigma(\mathfrak{B}^*\mathfrak{C}^*\mathfrak{C}\mathfrak{B}) \iff \lambda \in \sigma(\mathfrak{B}\mathfrak{B}^*\mathfrak{C}^*\mathfrak{C}) \\ &\iff \lambda \in \sigma(\mathcal{Q}\mathcal{O}) \iff \lambda \in \sigma(\mathcal{O}\mathcal{Q}). \end{aligned}$$

Since $\mathscr{X}$ is finite-dimensional in this instance the above equivalence can be strengthened to: for $\lambda \neq 0$

$$\lambda \in \sigma_p(H^*H) \iff \lambda \in \sigma_p(\mathcal{O}\mathcal{Q}),$$

as required. $\square$

Definition 2.1.6. The Lyapunov singular values $(\sigma_k)_{k=1}^m$ of a $H^\infty(\mathbb{C}_0^+, B(\mathscr{U}, \mathscr{Y}))$ transfer function are the singular values of its Hankel operator. They are always ordered such that $\sigma_k > \sigma_{k+1} > 0$, each with (geometric) multiplicity r_k.

Remark 2.1.7.

1. The Lyapunov singular values are sometimes simply called the singular values or Hankel singular values in the literature (for example in [66] and [34] respectively). We keep the term the Lyapunov singular value so as to distinguish them from the bounded real and positive real singular values introduced later.

2. Since the Lyapunov singular values are independent of the realisation chosen (because a Hankel operator and hence its singular values depend only on the transfer function, as will be explained in more detail in Section 5.1), it follows from Lemma 2.1.5 that the eigenvalues of the product $\mathcal{Q}\mathcal{O}$ are also independent of the realisation chosen.

3. When $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is in Lyapunov balanced co-ordinates it is easy to see that the Lyapunov singular values are precisely the eigenvalues of Π, where Π is as in (2.5).

We seek to outline how Lyapunov balanced realisations give rise to a natural truncation scheme and then describe how truncation (called Lyapunov balanced truncation) is performed. Suppose that $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is in Lyapunov balanced co-ordinates. Since Π is self-adjoint it is diagonalisable and so we can choose an eigenbasis $\{v_1, \ldots, v_n\}$ of Π of

$\mathscr{X}$. For each $j \in \{1, 2, \dots, n\}$, we let $\rho_j \in \{1, 2, \dots, m\}$ denote the integer such that

$$\Pi v_j = \sigma_{\rho_j} v_j. \tag{2.6}$$

Following the discussion on p. 6 and (2.6) we see that in Lyapunov balanced coordinates, states that are nearly uncontrollable are as equally nearly unobservable and vice versa. The concepts of controllability and observability are balanced. We also see that small singular values correspond to directions in the state-space that are nearly uncontrollable and so as equally nearly unobservable. Since the Lyapunov singular values are independent of the realisation chosen, the above remarks would suggest that eigenspaces corresponding to smaller singular values are a good candidate to omit in a reduced order system. Lyapunov balanced truncation is based on this principle.

Definition 2.1.8. Given rational $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ with Lyapunov singular values $(\sigma_k)_{k=1}^m$, let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal Lyapunov balanced realisation of G. For $r < m$, let $\mathscr{X}_r$ and $\mathscr{Z}_r$ denote the sum of the first r and last $m - r$ eigenspaces of Π respectively, with respective orthogonal projections $P_{\mathscr{X}_r}$ and $P_{\mathscr{Z}_r}$. Then with respect to the orthogonal decomposition $\mathscr{X} = \mathscr{X}_r \oplus \mathscr{Z}_r$, the operators A, B, C and Π split as

$$\begin{aligned}
\Pi &= \begin{bmatrix} P_{\mathscr{X}_r}\Pi|_{\mathscr{X}_r} & 0 \\ 0 & P_{\mathscr{Z}_r}\Pi|_{\mathscr{Z}_r} \end{bmatrix} = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{bmatrix}, & B &= \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \\
A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, & C &= \begin{bmatrix} C_1 & C_2 \end{bmatrix}.
\end{aligned}$$

The dimension of $\mathscr{X}_r$ is $\sum_{j=1}^r r_j$, the sum of the geometric multiplicities of the first r singular values. The truncated system with realisation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is called the reduced order system obtained by Lyapunov balanced truncation (of order $\sum_{j=1}^r r_j$), or simply the Lyapunov balanced truncation, and its transfer function is denoted by G_r.

The main result for Lyapunov balanced truncation is the following.

Theorem 2.1.9. *Given rational $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, let $(\sigma_k)_{k=1}^m$ denote the Lyapunov singular values and for $r < m$ let G_r denote the Lyapunov balanced truncation, then*

$$\|G - G_r\|_{H^\infty} \le 2 \sum_{j=r+1}^m \sigma_j. \tag{2.7}$$

If $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denotes a minimal, Lyapunov balanced realisation of G then the Lyapunov balanced truncation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is minimal and stable, and G_r has MacMillan degree $\sum_{j=1}^r r_j$.

Proof. For a proof of the error bound see Enns [23] or [26]. The remaining claims of stability, minimality (and hence degree) are proven in Pernebo and Silverman [66]. □

2.2 Bounded real balanced truncation

In this section we review model reduction by bounded real balanced truncation (for rational bounded real transfer functions). In the next section we do the same in the positive real case. As with Lyapunov balanced truncation, the aim is to obtain a reduced order transfer function by truncating an appropriate (i.e. balanced, in a sense to be made clear below) realisation, but to do so in such a way that the reduced order transfer function retains bounded realness or positive realness respectively.

We first recall the definition of bounded real.

Definition 2.2.1. Let $\mathscr{U}$ and $\mathscr{Y}$ denote Banach spaces. A function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is said to be bounded real if

$$\|G\|_{H^\infty} \leq 1. \tag{2.8}$$

We say that $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is strictly bounded real if

$$\|G\|_{H^\infty} < 1. \tag{2.9}$$

Remark 2.2.2.

1. We say that an input-state-output system is (strictly) bounded real if its transfer function is (strictly) bounded real.

2. The term bounded real is more common in the engineering than the mathematical literature, where Schur, contractive or scattering passive might be used instead. The same is true for positive real functions, which according to Staffans [79] (and the references therein), are also known as impedance passive functions, Caratheodory-Nevanlinna functions, Weyl functions or Titchmarsh-Weyl functions. Since we are considering model reduction by balanced truncation, where the terms bounded real and positive real are more common, we keep this convention.

3. Note also that there is no realness condition in Definition 2.2.1 (see also the definition of positive real, Definition 2.3.1).

2.2.1 The Bounded Real Lemma

Bounded real balanced truncation makes use of the well-known Bounded Real Lemma, see Anderson & Vongpanitlerd [2], which gives a state-space characterisation of bounded real functions. Since we will make frequent use of this result, we recall it below.

Proposition 2.2.3 (Bounded Real Lemma)**.** *Given $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, with $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ a minimal input-state-output realisation of G, the following are equivalent.*

(i) G is bounded real.

(ii) For input $u \in L^2(\mathbb{R}^+;\mathscr{U})$ and output $y \in L^2(\mathbb{R}^+;\mathscr{Y})$ with initial condition $x_0 = 0$

$$\int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \; ds \geq 0, \quad \forall\, t \geq 0.$$

(iii) There exists a positive, self-adjoint operator P on $\mathscr{X}$ such that for input $u \in L^2(\mathbb{R}^+;\mathscr{U})$ with output $y \in L^2(\mathbb{R}^+;\mathscr{Y})$ and initial state $x_0 \in \mathscr{X}$

$$\int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \; ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px_0, x_0\rangle_{\mathscr{X}}, \quad \forall\, t \geq 0.$$

(iv) There exists a triple of operators (P, K, W) with

$$P : \mathscr{X} \to \mathscr{X}, \quad K : \mathscr{X} \to \mathscr{U}, \quad W : \mathscr{U} \to \mathscr{U},$$

and P positive and self-adjoint satisfying the bounded real Lur'e equations

$$A^*P + PA + C^*C = -K^*K, \tag{2.10a}$$

$$PB + C^*D = -K^*W, \tag{2.10b}$$

$$I - D^*D = W^*W. \tag{2.10c}$$

The following are equivalent

(i)' G is strictly bounded real,

(ii)' There exists an $\varepsilon > 0$ such that for any input $u \in L^2(\mathbb{R}^+;\mathscr{U})$ and output $y \in L^2(\mathbb{R}^+;\mathscr{Y})$ with initial condition $x_0 = 0$

$$\int_0^t \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds \geq \varepsilon \int_0^t \|u(s)\|^2_{\mathscr{U}} + \|y(s)\|^2_{\mathscr{Y}} \, ds, \quad \forall\, t \geq 0.$$

(iv)' There exists a positive, self-adjoint operator P on $\mathscr{X}$, a solution of the so-called bounded real algebraic Riccati equation

$$A^*P + PA + C^*C + (PB + C^*D)(I - D^*D)^{-1}(B^*P + D^*C) = 0, \tag{2.11}$$

which is stabilising in the sense that

$$\sigma(A + B(I - D^*D)^{-1}(B^*P + D^*C)) \subseteq \mathbb{C}_0^-. \tag{2.12}$$

If any of (i) – (iv) hold then there are positive self-adjoint solutions P_m, P_M to (2.10) such that for any positive, self-adjoint solution P of (2.10) we have

$$0 < P_m \leq P \leq P_M. \tag{2.13}$$

The extremal operators P_m, P_M are the optimal cost operators of the bounded real optimal control problems, namely:

$$\langle P_M x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{u\in L^2(\mathbb{R}^-;\mathscr{U})\\ x(-\infty)=0,\, x(0)=x_0}} \int_{\mathbb{R}^-} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}}\, ds, \tag{2.14a}$$

$$-\langle P_m x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{u\in L^2(\mathbb{R}^+;\mathscr{U})\\ x(0)=x_0}} \int_{\mathbb{R}^+} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}}\, ds. \tag{2.14b}$$

The minimisation problems (2.14) *are subject to the minimal input-state-output realisation* (2.3)*, where $x(0) = x_0$ is the final state in* (2.14a) *and the initial state in* (2.14b)*. Similarly, if any of $(i)'$,$(ii)'$ or $(iv)'$ holds then there exists positive self-adjoint solutions P_m and P_M to* (2.11)*, extremal in the sense of* (2.13)*. Furthermore, P_m is stabilising in the sense of* (2.12) *and P_M is antistabilising in the sense that*

$$\sigma(A + B(I - D^*D)^{-1}(B^*P_M + D^*C)) \subseteq \mathbb{C}_0^+,$$

and P_m, P_M are the optimal cost operators as in (2.14)*.*

Proof. A proof of the equivalence of (i) and (iv) is given in [2]. The authors assume that $\dim \mathscr{U} = \dim \mathscr{Y}$, but the result is true in general. A short series of calculations gives the implications $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$. A proof of the equivalence of $(i)'$ and $(ii)'$ can be found in Zhou *et al.* [104, Theorem 13.19 and Corollary 13.24]. Another short series of calculations gives $(i)' \iff (ii)'$ □

If $P = P^* > 0$ is a solution of (2.10), for some K, W then an elementary calculation shows that $P^{-1} > 0$ solves the dual bounded real Lur'e equations,

$$AQ + QA^* + BB^* = -LL^*, \tag{2.15a}$$

$$QC^* + BD^* = -LX^*, \tag{2.15b}$$

$$I - DD^* = XX^*, \tag{2.15c}$$

for some operators $L : \mathscr{Y} \to \mathscr{X}$, $X : \mathscr{Y} \to \mathscr{Y}$. By the Bounded Real Lemma, there are extremal self-adjoint solutions Q_m, Q_M to (2.15) such that for any self-adjoint solution Q to (2.15); $0 < Q_m \le Q \le Q_M$. These observations are another way of expressing that G is bounded real if and only if the dual transfer function

$$\mathbb{C}_0^+ \ni s \mapsto G_d(s) = [G(\bar{s})]^*,$$

is bounded real. In particular, it is not difficult to see that

$$P_m = Q_M^{-1}, \quad \text{and} \quad P_M = Q_m^{-1}. \tag{2.16}$$

Remark 2.2.4. Solutions of the bounded real Lur'e equations (2.10) are generally not unique. Moreover, given solutions (P, K, W) and (Q, L, X) of the bounded real Lur'e equations (2.10) and dual bounded real Lur'e equations (2.15) respectively, the first components P and Q do not in general uniquely determine the other two respective components. We expand on this further, as there are important parallels to the infinite-dimensional case of Chapter 6. If we fix (P, K, W) and (Q, L, X) as above then the operators K', W' and L', X' defined by

$$\begin{aligned} K' &= UK, \quad W' = UW \\ L' &= LV, \quad\;\; X' = XV, \end{aligned}$$

for $U : \mathscr{U} \to \mathscr{U}$, $V : \mathscr{Y} \to \mathscr{Y}$ unitary, are such that (P, K', W') and (Q, L', X') are also solutions of (2.10) and (2.15) respectively. If G is strictly bounded real then this is all the freedom there is. In the non-strict case there is more freedom. If $I - D^*D$ is invertible then we can obtain a special solution by taking W as the non-negative, invertible, square root of $I - D^*D$, which is unique, and also determines K via (2.10b). Similarly for X and L.

2.2.2 Bounded real balanced truncation

Bounded real balanced truncation is very similar in principle to Lyapunov balanced truncation only now the quantities to be balanced are the self-adjoint, positive optimal cost operators P_M and P_m from (2.14).

Definition 2.2.5. We say that the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is bounded real balanced, or in bounded real balanced co-ordinates, if

$$P_m = P_M^{-1} =: \Pi. \tag{2.17}$$

The non-negative square roots of the eigenvalues of the product $P_m P_M^{-1}$ are called the bounded real singular values, which we denote by $(\sigma_k)_{k=1}^m$, each with (geometric) multiplicity r_k, (so that $\sum_{k=1}^m r_k = \dim \mathscr{X}$). The bounded real singular values are ordered such that $\sigma_k > \sigma_{k+1} > 0$ for each k.

Remark 2.2.6.

1. From condition (2.13) we see that every bounded real singular value σ_k satisfies $\sigma_k \in (0, 1]$ for all $k \in \{1, 2 \dots, m\}$. Furthermore, equality (2.16) implies that the bounded real singular values are the square roots of the eigenvalues of $P_m Q_m$. In practise it is sometimes easier to compute Q_m than to compute P_M^{-1}.
2. As with the Lyapunov singular values, the bounded real singular values are independent of the realisation of the bounded real transfer function chosen. This is proven in Lemma 6.3.12.

3. Given a minimal stable realisation of a bounded real transfer function there always exists a similarity transformation such that the transformed realisation is bounded real balanced, for the same reasons as in the Lyapunov balanced case.

4. If a realisation is bounded real balanced then the bounded real singular values are precisely the eigenvalues of Π as in (2.17).

We describe the motivation for bounded real balanced truncation. By the Bounded Real Lemma we see that the right hand side of (2.14a) is a function of the final state x_0 and is the minimal energy (with respect to the bounded real supply rate) required to reach x_0 from zero in infinite time. Similarly, the right hand side of (2.14b) is a function of the initial state x_0 and the maximal energy (again with respect to the bounded real supply rate) we can extract from the state x_0 in infinite time. We denote these quantities by C_{x_0} and E_{x_0} respectively. Suppose that $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is bounded real balanced, and let $\{v_1, \dots, v_n\}$ denote a basis of $\mathscr{X}$ of eigenvectors of Π. If for each $j \in \{1, 2, \dots, n\}$, σ_{ρ_j} denotes the eigenvalue of Π corresponding to v_j for every j then from (2.14) and (2.17) we see that

$$E_{v_j} = -\sigma_{\rho_j},\ C_{v_j} = \sigma_{\rho_j}^{-1} \quad \text{and so} \quad E_{v_j} = -C_{v_j}^{-1}.$$

Here the minus sign reflects the difference of energy flow into or out of the system. We draw the same conclusions as in the Lyapunov balanced case, namely that if $\sigma_{\rho_j} = 1$ then the energy required to reach the state v_j is equal to the energy we can extract from that state. Additionally, the energy required to reach state v_j is inversely proportional to the energy obtained from v_j. Bounded real balanced truncation is therefore based on truncating according to the size of the bounded real singular values.

Bounded real balanced truncation is performed in an identical manner to Lyapunov balanced truncation, described in Definition 2.1.8 only now for $r < m$, $\mathscr{X}_r$ and $\mathscr{Z}_r$ denote the sum of the first r and last $m-r$ eigenspaces of Π given by (2.17) respectively. The truncated system with realisation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is called the reduced order system obtained by bounded real balanced truncation (of order $\sum_{j=1}^{r} r_j$, the sum of the geometric multiplicities of the first r bounded real singular values), or just the bounded real balanced truncation. The transfer function of $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is called the reduced order transfer function obtained by bounded real balanced truncation and is denoted by G_r.

The main result for bounded real balanced truncation for rational transfer functions is stated below.

Theorem 2.2.7. *Given $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ bounded real, let $(\sigma_j)_{j=1}^m$ denote the bounded real singular values, with multiplicities r_j. For $r < m$ let G_r denote the reduced order transfer obtained by bounded real balanced truncation. Then G_r is bounded real* *and the following error bound holds*

$$\|G - G_r\|_{H^\infty} \leq 2 \sum_{j=r+1}^{m} \sigma_j. \tag{2.18}$$

Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *denote a minimal, bounded real balanced realisation of* G*. Then the bounded real balanced truncation* $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ *is stable. If additionally* G *is strictly bounded real, then* G_r *has MacMillan degree* $\sum_{j=1}^{r} r_j$ *and* $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ *is minimal and bounded real balanced.*

Proof. See Theorem 2 and Section IV of [57]. The assumption there that G is strictly bounded real is not needed to prove that G_r is bounded real, that A_{11} is stable and that the error bound (2.18) holds. The authors also assume throughout that $\mathscr{U} = \mathscr{Y}$, but this isn't needed and the proof for the general case is essentially the same. □

2.2.3 Relation to Lyapunov balanced truncation

Given a bounded real balanced realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, which we denote by Σ, observe that from (2.10a), the optimal cost operator $\Pi = P_m$ satisfies the Lyapunov equation

$$A^*\Pi + \Pi A + \begin{bmatrix} C^* & K^* \end{bmatrix} \begin{bmatrix} C \\ K \end{bmatrix} = 0, \tag{2.19}$$

which in the strict case also follows from the Riccati equation (2.11) by setting

$$K := (I - D^*D)^{-\frac{1}{2}}(B^*\Pi + C^*D).$$

Similarly from (2.15a), $\Pi = Q_m$ satisfies the Lyapunov equation

$$A\Pi + \Pi A^* + \begin{bmatrix} B & L \end{bmatrix} \begin{bmatrix} B^* \\ L^* \end{bmatrix} = 0. \tag{2.20}$$

Since A is stable it follows that Π is the controllability and observability Gramians of the extended system

$$\left[\begin{array}{c|cc} A & B & L \\ \hline C & D & X \\ K & W & 0 \end{array}\right], \tag{2.21}$$

which we denote by Σ_E. Note that Σ_E itself depends on Π through K and L and also by Remark 2.2.4, Σ_E is not uniquely determined by A, B, C, D and Π. For every choice of K, L, X and W, however, Σ_E has input and output spaces $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$ respectively, and the same state-space as Σ.

From the Lyapunov equations (2.19) and (2.20) we see that Σ_E is Lyapunov balanced and by Lemma 2.1.5 the bounded real singular values of Σ are the Lyapunov singular values of Σ_E. The Lyapunov balanced truncation of (2.21) (with the same order as the bounded real balanced truncation) is

$$\left[\begin{array}{c|cc} A_{11} & B_1 & L_1 \\ \hline C_1 & D & X \\ K_1 & W & 0 \end{array}\right],$$

from which the bounded real balanced truncation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D_1 \end{smallmatrix}\right]$ of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is recovered by omitting the blocks L_1, K_1, X, W and zero. This corresponds to restricting to and projecting onto the original input and output spaces $\mathscr{U}$ and $\mathscr{Y}$ respectively. Therefore bounded real balanced truncation of Σ can be seen as *Lyapunov* balanced truncation of Σ_E and moreover although Σ_E is not unique, the bounded real balanced truncation is. This relation is used in [57] in proving the bounded real balanced truncation error bound (2.18) from Theorem 2.2.7.

2.3 Positive real balanced truncation

We recall the definition of positive real.

Definition 2.3.1. An operator valued analytic function $J : \mathbb{C}_0^+ \to B(\mathscr{U})$, where $\mathscr{U}$ is a Hilbert space, is positive real if

$$J(s) + [J(s)]^* \geq 0, \quad \forall\, s \in \mathbb{C}_0^+. \tag{2.22}$$

We say that the analytic function $J : \mathbb{C}_0^+ \to B(\mathscr{U})$ is strictly positive real if there exists $\eta > 0$ such that

$$J(s) + [J(s)]^* \geq \eta I, \quad \forall\, s \in \mathbb{C}_0^+. \tag{2.23}$$

Remark 2.3.2.

1. We say that an input-state-output system is (strictly) positive real if its transfer function is (strictly) positive real. Observe that positive real systems are "square", in so much that the input and output spaces are the same.

2. The requirement that J is analytic in Definition 2.3.1 is crucial. When defining positive real for rational functions it is sometimes omitted (since clearly rational functions are certainly analytic away from their poles).

3. The term strictly positive real is used for various slightly different concepts in the literature. Some of these differences are explained in, for example, Wen [98]. The condition (2.23) is equivalent to the concept called extended strictly positive real, used in for example Sun *et al.* [82, Definition 2.1].

4. We do not assume that a positive real function is real on the real axis. There is no mathematical advantage to making such an assumption, and hence we have omitted it.

2.3.1 The Positive Real Lemma

Positive real balanced truncation makes use of the Positive Real Lemma, a famous result from the 1960s, which is often known as the KYP Lemma after the contributions of Kalman, Yakubovic and Popov, which gives a state-space characterisation of positive real functions. See [2, p. 218] for more details of the history. As we will make use of this result we recall it here.

Proposition 2.3.3 (Positive Real Lemma)**.** *Given $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, with $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ a minimal input-state-output realisation of J, the following are equivalent.*

(i) J is positive real.

(ii) For input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ and output $y \in L^2(\mathbb{R}^+; \mathscr{U})$ with initial condition $x_0 = 0$

$$\int_0^t 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\, ds \geq 0, \quad \forall\, t \geq 0.$$

(iii) There exists a positive, self-adjoint operator P on $\mathscr{X}$ such that for input $u \in L^2(\mathbb{R}^+; \mathscr{U})$ with output $y \in L^2(\mathbb{R}^+; \mathscr{U})$ and initial state $x_0 \in \mathscr{X}$

$$\int_0^t 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\, ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px_0, x_0\rangle_{\mathscr{X}}, \quad \forall\, t \geq 0.$$

(iv) There exists a triple of operators (P, K, W) with

$$P : \mathscr{X} \to \mathscr{X}, \quad K : \mathscr{X} \to \mathscr{U}, \quad W : \mathscr{U} \to \mathscr{U},$$

and P positive and self-adjoint satisfying the positive real Lur'e equations

$$A^*P + PA = -K^*K, \tag{2.24a}$$

$$PB - C^* = -K^*W, \tag{2.24b}$$

$$D + D^* = W^*W. \tag{2.24c}$$

The following are equivalent.

(i)' J is strictly positive real.

(iv)' There exists a positive, self-adjoint operator P on $\mathscr{X}$ satisfying the positive real

algebraic Riccati equation

$$A^*P + PA + (PB - C^*)(D + D^*)^{-1}(B^*P - C) = 0. \tag{2.25}$$

which is stabilising in the sense that

$$\sigma(A + B(D + D^*)^{-1}(B^*P - C)) \subseteq \mathbb{C}_0^-. \tag{2.26}$$

If any of $(i) - (iv)$ *hold then there are positive, self-adjoint solutions* $\tilde{P}_m, \tilde{P}_M$ *to* (2.24) *such that any positive, self-adjoint solution* P *to* (2.24) *satisfies*

$$0 < \tilde{P}_m \leq P \leq \tilde{P}_M. \tag{2.27}$$

The extremal operators $\tilde{P}_m$*,* $\tilde{P}_M$ *are the optimal cost operators of the positive real optimal control problems, namely:*

$$\langle \tilde{P}_M x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{u \in L^2(\mathbb{R}^-;\mathscr{U}) \\ x(-\infty)=0,\, x(0)=x_0}} \int_{\mathbb{R}^-} 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\, ds, \tag{2.28a}$$

$$-\langle \tilde{P}_m x_0, x_0\rangle_{\mathscr{X}} = \inf_{\substack{u \in L^2(\mathbb{R}^+;\mathscr{U}) \\ x(0)=x_0}} \int_{\mathbb{R}^+} 2\,\mathrm{Re}\,\langle u(s), y(s)\rangle_{\mathscr{U}}\, ds. \tag{2.28b}$$

The minimisation problems (2.28) *are subject to the minimal input-state-output realisation* (2.3)*, where* $x(0) = x_0$ *is the final state in* (2.28a) *and the initial state in* (2.28b)*. Similarly, if* $(i)'$ *or* $(ii)'$ *holds then there exists positive self-adjoint solutions* $\tilde{P}_m$ *and* $\tilde{P}_M$ *to* (2.25)*, extremal in the sense of* (2.27)*. Furthermore,* $\tilde{P}_m$ *is stabilising in the sense of* (2.26) *and* $\tilde{P}_M$ *is antistabilising in the sense that*

$$\sigma(A + B(D + D^*)^{-1}(B^*\tilde{P}_M - C)) \subseteq \mathbb{C}_0^+,$$

and $\tilde{P}_m$*,* $\tilde{P}_M$ *are the optimal cost operators as in* (2.28)*.*

Proof. A proof of the equivalence of (i) and (iv) is given in Section 5.2 of [2]. For the equivalence of (i) and (ii) see Willems [99, Theorem 1]. A short series of calculations shows $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$. A proof of $(i)' \iff (ii)'$ is given in [104, Corollary 13.27]. $\square$

An elementary calculation demonstrates that if $P = P^* > 0$ solves (2.24), for some

K, W, then P^{-1} solves the dual positive real Lur'e equations

$$AQ + QA^* = -LL^*, \tag{2.29a}$$

$$QC^* - B = -LX^*, \tag{2.29b}$$

$$D + D^* = XX^*, \tag{2.29c}$$

for some operators $L : \mathscr{U} \to \mathscr{X}$, $X : \mathscr{U} \to \mathscr{U}$.

By the Positive Real Lemma, there are positive self-adjoint solutions $\tilde{Q}_m$, $\tilde{Q}_M$ to (2.29) such that for any self-adjoint solution Q to (2.29) it follows that $0 < \tilde{Q}_m \leq Q \leq \tilde{Q}_M$. We see that J is positive real if and only if the dual transfer function J_d given by

$$J_d(s) = [J(\overline{s})]^*,$$

is. Furthermore, it readily follows that

$$\tilde{P}_m = \tilde{Q}_M^{-1}, \quad \text{and} \quad \tilde{P}_M = \tilde{Q}_m^{-1}. \tag{2.30}$$

2.3.2 Positive real balanced truncation

Positive real balanced truncation is identical in spirit to bounded real balanced truncation and makes use of the optimal cost operators $\tilde{P}_m$ and $\tilde{P}_M$ from (2.28).

Definition 2.3.4. We say that the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is positive real balanced, or in positive real balanced co-ordinates, if

$$\tilde{P}_m = \tilde{P}_M^{-1} =: \Pi. \tag{2.31}$$

The non-negative square roots of the eigenvalues of the product $\tilde{P}_m \tilde{P}_M^{-1}$ are called the positive real singular values, which we denote by $(\sigma_k)_{k=1}^m$, each with (geometric) multiplicity r_k, (so that $\sum_{k=1}^m r_k = \dim \mathscr{X}$). The positive real singular values are ordered such that $\sigma_k > \sigma_{k+1} > 0$ for each k.

The positive real balanced truncation is defined in the same way as the bounded real balanced truncation. Note from (2.30) that the positive real singular values are the (non-negative) square roots of the eigenvalues of $\tilde{P}_m \tilde{Q}_m$. The main result for positive real balanced truncation is stated below.

Theorem 2.3.5. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real transfer function and let $(\sigma_j)_{j=1}^m$ denote the positive real singular values, each with multiplicity r_j. For $r < m$, let J_r denote the reduced order transfer obtained by positive real balanced truncation. Then $J_r \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ and J_r is positive real. If $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denotes the minimal positive real balanced realisation of J then the positive real balanced truncation $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$*

is stable. If additionally J is strictly positive real, then J_r has MacMillan degree $\sum_{j=1}^{r} r_j$ and $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$ is minimal and positive real balanced.

Proof. See Harshavardhana *et al.* [40] and the references therein. □

The next result contains error bounds for positive real balanced truncation.

Theorem 2.3.6. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function with minimal realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and let $(\sigma_j)_{j=1}^m$ denote the positive real singular values, each with multiplicity r_j. For $r < m$, let J_r denote the reduced order transfer obtained by positive real balanced truncation. Then the following bounds hold*

(i) $\left\|(D^* + J)^{-1} - (D^* + J_r)^{-1}\right\|_{H^\infty} \leq 2\|(D + D^*)^{-1}\| \sum_{j=r+1}^{m} \sigma_j$,

(ii) $\left\|(D^* + J)^{-1}[J - J_r](D^* + J_r)^{-1}\right\|_{H^\infty} \leq 2\|(D + D^*)^{-1}\| \sum_{j=r+1}^{m} \sigma_j$,

(iii) $\left\|(D^* + J_r)^{-1}(J - J_r)\right\|_{H^\infty} \leq 2\|(D + D^*)^{-1}\| \, \|D^* + J\|_{H^\infty} \sum_{j=r+1}^{m} \sigma_j$.

Proof. See [3, Proposition 7.17] (or [34, Theorem 5]) for a proof of (i). The bound in (ii) is equivalent to that in (i), see [3, Remark 7.5.2]. For a proof of (iii) see [34, Lemma 3]. □

Remark 2.3.7. The H^∞ error bound (2.18) for bounded real balanced truncation follows from the corresponding error bound (2.7) for Lyapunov balanced truncation. Bounded real balanced truncation can be thought of Lyapunov balanced truncation of a certain extended system as described in Section 2.2.3. In the positive real case, however, there is *not* the same relationship to Lyapunov balanced truncation. This is because of the different structure of the positive real equations (2.24). The following example demonstrates that the analogous H^∞ error bound *does not hold* in the positive real case. Consider

$$\mathbb{C}_0^+ \ni s \mapsto J(s) = 1 + \frac{s}{s+1} = 2 - \frac{1}{s+1}.$$

It is easy to see that J is positive real as

$$J(s) + [J(s)]^* = 2\,\mathrm{Re}\,J(s) = 2\left[2 - \mathrm{Re}\left(\frac{s+1}{|s+1|^2}\right)\right] \geq 0, \quad \forall s \in \mathbb{C}_0^+,$$

and that

$$A = -1, \quad B = 1, \quad C = -1, \quad D = 2,$$

is a (minimal) realisation of J. In this instance as $D + D^*$ is invertible, the positive real Lur'e equations (2.24) collapse to the positive real algebraic Riccati equation (2.25), which is a scalar quadratic equation with extremal solutions

$$\tilde{P}_m = 3 - 2\sqrt{2}, \quad \tilde{P}_M = 3 + 2\sqrt{2}.$$

The positive real singular value is $\sigma = \tilde{P}_m \tilde{P}_M^{-1} = 17 - 12\sqrt{2} = 0.0294$. Thus for $r = 0$ we have $J_r = D$ and as

$$|J(0) - D| = 1 > 2\sigma,$$

the H^∞ error bound cannot hold.

In fact in the positive real case an H^∞ error bound seems less natural, and in Corollary 3.6.9 we obtain the following gap metric error bound for positive real balanced truncation. Namely,

$$\hat{\delta}(J, J_r) \leq 2 \sum_{k=r+1}^{m} \sigma_k, \tag{2.32}$$

where σ_k are the positive real singular values and $\hat{\delta}$ is the gap metric. The gap metric is described in more detail in Section 3.1.4.

2.3.3 A false error bound in the literature

The error bound for positive real balanced truncation

$$\|J - J_r\|_{H^\infty} \leq \|D + D^*\| \sum_{j=r+1}^{m} \frac{2\sigma_j}{(1-\sigma_j)^2} \left(1 + \sum_{i=1}^{j-1} \frac{2\sigma_i}{1-\sigma_i}\right), \tag{2.33}$$

claimed in [14, Theorem 2] is false, as the following counter-example demonstrates.

Consider the following continuous time, time invariant single input, single output linear system on the state-space $\mathbb{C}^4$:

$$\begin{aligned} M\dot{\mathtt{x}}(t) &= K\mathtt{x}(t) + L\mathtt{u}(t), \\ \mathtt{y}(t) &= H\mathtt{x}(t) + D\mathtt{u}(t), \end{aligned} \tag{2.34}$$

where

$$\begin{aligned} M &= \begin{bmatrix} \frac{1}{12} & \frac{1}{24} & 0 & 0 \\ \frac{1}{24} & \frac{1}{6} & \frac{1}{24} & 0 \\ 0 & \frac{1}{24} & \frac{1}{6} & \frac{1}{24} \\ 0 & 0 & \frac{1}{24} & \frac{1}{6} \end{bmatrix}, \qquad L = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \\ K &= \begin{bmatrix} -4 & 4 & 0 & 0 \\ 4 & -8 & 4 & 0 \\ 0 & 4 & -8 & 4 \\ 0 & 0 & 4 & -8 \end{bmatrix}, \quad H = L^*, \\ D &= 0.01. \end{aligned} \tag{2.35}$$

The physical motivation for studying (2.34) comes from a finite element approximation of the heat equation

$$\left.\begin{aligned} w_t(t,x) &= w_{xx}(t,x), \\ w(0,x) &= w_0(x), \\ w(t,1) &= 0, \end{aligned}\right\} \quad t \geq 0,\ x \in [0,1], \tag{2.36}$$

with input $\mathtt{u}$ and output $\mathtt{y}$ satisfying

$$\begin{aligned} \mathtt{u}(t) &:= w_x(t,0), \\ \mathtt{y}(t) &:= -w(t,0) + Dw_x(t,0). \end{aligned} \tag{2.37}$$

By setting $A := M^{-1}K$, $G = M^{-1}L$, we can rewrite (2.34) as

$$\begin{aligned} \dot{\mathtt{x}}(t) &= A\mathtt{x}(t) + G\mathtt{u}(t), \\ \mathtt{y}(t) &= H\mathtt{x}(t) + D\mathtt{u}(t), \end{aligned} \tag{2.38}$$

with transfer function

$$J(s) = D + H(sI - A)^{-1}G. \tag{2.39}$$

Observe that the system with transfer function $J - D$ is positive real as $P = M = P^* > 0$, $N = \sqrt{-2K}$ and $R = 0$ satisfy the positive real Lur'e equations (2.24). Therefore for $s \in \mathbb{C}$ with $\operatorname{Re} s \geq 0$,

$$\begin{aligned} &[(J - D)(s)]^* + (J - D)(s) \geq 0, \\ &\qquad \Rightarrow [Z(s)]^* + Z(s) \geq 2D > 0, \end{aligned}$$

and so the system (2.38) is strictly positive real (extended strictly positive real in [14]). It is easy to verify also that (2.39) is a minimal, and hence controllable and observable, realisation of J. The positive real singular values of Σ are

$$\sigma_1 = 0.6640,\ \sigma_2 = 0.2927,\ \sigma_3 = 0.0487,\ \sigma_4 = 0.0036. \tag{2.40}$$

The first order positive real balanced truncation of Σ is

$$\hat{J}(s) = \frac{0.01s + 12.74}{s + 51.97},$$

and the approximation error $\|J - \hat{J}\|_{\mathcal{H}^\infty}$ is 0.7648. However, the error bound provided in [14, Theorem 2] is

$$2D \sum_{i=2}^{4} \frac{2\sigma_i}{(1-\sigma_i)^2} \left(1 + \sum_{j=1}^{i-1} \frac{2\sigma_j}{1-\sigma_j}\right)^2 = 0.6509,$$

which is smaller than the error. Hence [14, Theorem 2] is false.

Remark 2.3.8. We remark that there is some confusion in the literature regarding the nomenclature balanced stochastic truncation (the term that was used in [14]). Originally balanced stochastic truncation of a positive real function J meant a reduced order triple J_r, V_r, W_r with V_r and W_r left and right spectral factors of $J_r + J_r^*$ respectively, which are obtained by balancing the minimal nonnegative definite solutions of the (primal and dual) positive real equations and truncating. Nowadays ([3, p. 229] or [34]) the term positive real balanced truncation is used for obtaining only J_r in this way, and the term balanced stochastic truncation is reserved for a generalization of obtaining V_r from a function V which can be seen as a left spectral factor of $J + J^*$. The matlab function `bstmr` (balanced stochastic truncation model reduction) for example only does the latter. The article [14] however pertains to what is now called positive real balanced truncation.

The proof of [14, Theorem 2] fails because for our above example the bound (18) in [14] is false. Using the notation of [14] (note here only one state is truncated from Σ) it follows that

$$\|T_1\|_\infty = 4.0389 > 1.7692 = 2\sum_{i=1}^{3} \frac{\sigma_i^2}{1-\sigma_i^2}. \tag{2.41}$$

Their proof of bound (18) uses [90, Lemma 5], which is only proven in [90] under the assumptions (51) and (53) (using the numbering of [90]). However, the authors state that [90, Lemma 5] also holds when (51) and (54) are satisfied. The above example shows that this is false. Letting

$$S = T_1,\ P(s) = Q(s) = \operatorname{diag}(\sigma_1, \sigma_2, \sigma_3) =: \hat{\Pi},$$

then equations (51) and (54) from [90] hold with A, B and C replaced by $\hat{A}_1, \hat{B}_1$ and $\hat{C}_1$ (again, notation from [14]), but the conclusion fails as inequality (2.41) shows. In this instance,

$$\hat{A}_1^*\hat{\Pi} + \hat{\Pi}\hat{A}_1 + \hat{C}_1^*\hat{C}_1 \neq 0,$$

and so equation (53) of [90] does not hold.

Counter-examples to [90, Theorem 1], which also uses the flawed [90, Lemma 5] in its proof, can be found in Chen & Zhou [13] and Zhou *et al.* [104, p. 171]. It is not pointed out there, however, that a flaw to [90, Theorem 1] occurs in [90, Lemma 5].

2.4 Notes

Lyapunov balanced truncation was introduced by Moore in [52] as a means of model reduction. The H^∞ error bound (2.7) which justifies this was obtained independently by Enns [23] and Glover [26]. Lyapunov balanced truncation can only be applied to stable systems, the variant LQG-balanced truncation introduced by Verriest [89] and studied further by among others Jonckheere & Silverman [41] overcomes this problem. LQG-balanced truncation is equivalent to Lyapunov balanced truncation of normalized coprime factors [48], [49]. This connection leads to an error bound in the gap metric which justifies LQG balanced truncation as an approximation method.

As a testimony to its efficacy, there have been many extensions to the original Lyapunov balanced truncation. For example, as mentioned in the introduction, to positive real balanced truncation, which was considered first by Desai & Pal in [22], and has been contributed to by Harshavardhana *et al.* [39], [40], and also to bounded real balanced truncation, introduced by Opdenacker & Jonckheere in [57]. Other extensions include to balanced stochastic truncation in Green [32] and to behavioral systems in Minh [51].

Almost everything in this chapter is already known, and the survey article by Gugercin & Antoulas [34], as well as Antoulas [3] include summaries of the material. The counter-example to the error-bound (2.33) from [14] is new, and has been published in [35]. We are grateful to Timo Reis for his suggestions with this matter. He first suggested that the result of [14] may be false because it relied on [90]. He also directed our attention to the counter-examples of [90, Theorem 1].

Chapter 3

Dissipative balanced truncations

In this chapter we seek to generalise and also unify bounded real and positive real balanced truncation described in Chapter 2. We restrict attention to the finite-dimensional case. It is well known that bounded real and positive real input-state-output systems are related by the Cayley transform, also known as the diagonal or Möbius transform. Our approach, however, is to consider dissipative state-space systems, which we recall in the chapter and which are a behavioral object in the sense that they make no distinction between inputs and outputs. Instead, these systems contain an external signal that incorporates interactions with the surrounding environment and which we demonstrate can often be decomposed into an input and output so that classical input-state-output systems are recovered. In such a framework we show that bounded real and positive real input-state-output systems are particular manifestations of a dissipative state-signal system. Accordingly, we derive dissipative balanced truncation for dissipative state-signal systems and the main results of this chapter are a gap metric error bound for dissipative balanced truncation, formulated as Theorem 3.6.8. Also of interest are a "generalised" KYP Lemma, Theorem 3.5.5, of which the Bounded Real and Positive Real Lemmas in Chapter 2 can be seen as special cases.

The motivation for such an approach is twofold; firstly, in some instances the distinction in a mathematical model between inputs and outputs is unclear and model reduction by balanced truncation in an environment free from those constrictions is desirable. Secondly, the gap metric error bound, which is not an input-output object, supports this as a model reduction scheme. As a corollary we also obtain a gap metric error bound for positive real balanced truncation, Corollary 3.6.9.

3.1 State space systems

In this section we define precisely what we mean by an input-state-output system, a driving-variable system and an output-nulling system. These are examples of state space systems as in Willems [100] or state/signal systems as studied in the discrete time infinite-dimensional case by Arov & Staffans [4]-[8]. More recently state/signal systems have been studied in continuous time by Staffans & Kurula [44], [45]. These objects and the relations between them are a key ingredient of this chapter.

3.1.1 Definitions

We begin with a remark on what we mean by the solution of an ODE.

Remark 3.1.1. Let $\mathscr{U}, \mathscr{X}$ denote finite-dimensional Hilbert spaces and let A, B denote operators

$$A : \mathscr{X} \to \mathscr{X}, \quad B : \mathscr{U} \to \mathscr{X}.$$

For $u \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{U})$ by a solution x of the (formal) ordinary differential equation

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ x(0) &= x_0, \end{aligned} \quad t \geq 0, \tag{3.1}$$

we mean the continuous function $x \in C(\mathbb{R}^+; \mathscr{X})$ given by the variation of parameters formula

$$\mathbb{R}^+ \ni t \mapsto x(t) = \mathrm{e}^{At} x_0 + \int_0^t \mathrm{e}^{A(t-s)} Bu(s)\, ds. \tag{3.2}$$

In the above e^A denotes the matrix exponential of A. In fact, x given by (3.2) belongs to the Sobolev space $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{X})$ and thus the equation (3.1) holds for almost all $t \geq 0$.

Definition 3.1.2. Given $\mathscr{U}, \mathscr{X}$ and $\mathscr{Y}$ finite-dimensional Hilbert spaces we define an input-state-output node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{3.3}$$

with associated formal differential equation

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad x(0) = x_0. \tag{3.4}$$

The spaces $\mathscr{U}, \mathscr{X}$ and $\mathscr{Y}$ are called the input, state and output spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\mathrm{loc}}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) \end{bmatrix} : \exists\, x_0 \in \mathscr{X} \text{ such that (3.4) holds} \right\}. \tag{3.5}$$

The component x of a trajectory is understood as a solution of (3.4) as described in

Remark 3.1.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T} \;:\; x(0) = x_0 \right\}, \tag{3.6}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\begin{aligned} \mathcal{T}_{\text{ext}} := \Bigg\{ \begin{bmatrix} u \\ y \end{bmatrix} \in L^2_{\text{loc}}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \\ \text{such that} \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}(0) \Bigg\}. \end{aligned} \tag{3.7}$$

We call the pair consisting of the node (3.3) and set of trajectories (3.5) an input-state-output system, which we denote by $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$.

Definition 3.1.3. Given an input-state-output system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and set of trajectories $\mathcal{T}$ we define the transfer function G as the operator valued function of a complex variable

$$\rho(A) \ni s \mapsto G(s) = D + C(sI - A)^{-1}B,$$

where $\rho(A)$ is the resolvent set of A. We define the input-output map $\mathfrak{D} : L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{U}) \to L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{Y})$ as

$$y = \mathfrak{D}u, \qquad \text{where } u, y \text{ are such that } \begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{T}_{\text{ext}}. \tag{3.8}$$

Definition 3.1.4. Given an input-state-output system, with set of trajectories $\mathcal{T}$, we define the set of stable externally generated trajectories $\mathcal{S}$ as

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]).$$

We say that the input-state-output system is (input-output) stable if the projection of $\mathcal{S}$ onto $L^2(\mathbb{R}^+; \mathscr{U})$ is all of $L^2(\mathbb{R}^+; \mathscr{U})$.

Proposition 3.1.5. *The input-output map $\mathfrak{D}$ of a stable input-state-output system is bounded $L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$. If $\mathcal{S}$ denotes the set of stable externally generated trajectories then*

$$\mathcal{S} = \mathcal{G}(\mathfrak{D}). \tag{3.9}$$

Moreover, the externally generated trajectories $\mathcal{T}_{\text{ext}}$ are characterised by $\mathfrak{D}$ through (3.8) *and in the stable case* (3.9). *The input-output map $\mathfrak{D}$ and transfer function G uniquely* *determine one another via the relationship*

$$G(s)\hat{u}(s) = \widehat{\mathfrak{D}u}(s), \tag{3.10}$$

where $\hat{f}$ *is the Laplace transform of* f*. Therefore,* $\mathcal{T}_{\text{ext}}$ *is characterised by* G *through* (3.8)*,* (3.9) *and* (3.10)*.*

Proof. This is standard in input-state-output systems theory. □

Definition 3.1.6. Given $\mathscr{V}, \mathscr{X}$ and $\mathscr{W}$ finite-dimensional Hilbert spaces we define a driving-variable node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix}, \tag{3.11}$$

where $D : \mathscr{V} \to \mathscr{W}$ is assumed injective, with associated formal differential equation

$$\begin{bmatrix} \dot{x}(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix}, \quad x(0) = x_0. \tag{3.12}$$

The spaces $\mathscr{V}, \mathscr{X}$ and $\mathscr{W}$ are called the driving-variable, state and signal spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) \end{bmatrix} : \exists\, x_0 \in \mathscr{X},\ v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V}) \right. \\ \left. \text{such that (3.12) holds} \right\}. \tag{3.13}$$

The component x of a trajectory is understood as a solution of (3.12) as described in Remark 3.1.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T} \ : \ x(0) = x_0 \right\}, \tag{3.14}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\mathcal{T}_{\text{ext}} := \left\{ w \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \text{ such that } \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T}(0) \right\}. \tag{3.15}$$

We define the set of stable externally generated trajectories $\mathcal{S}$ by

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \mathscr{W}). \tag{3.16}$$

We call the pair consisting of the node (3.11) and set of trajectories (3.13) a driving-variable system, which we denote by $\left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T}\right)_{\text{dv}}$.

Remark 3.1.7. Although a driving-variable system looks like a standard input-state-output system, its interpretation is very different. The external signal w incorporates all the interaction with the external world (so in the standard input-state-output formulation it would contain both the outputs *and* the inputs). The driving-variable v is a latent variable used to mathematically describe the dynamics and may or may not have any physical meaning (much like a state).

Definition 3.1.8. Given $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ finite-dimensional Hilbert spaces we define an output-nulling node as an operator

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix}, \tag{3.17}$$

where $D : \mathscr{W} \to \mathscr{V}$ is assumed surjective, with associated formal algebraic-differential equation

$$\begin{bmatrix} \dot{x}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix}, \quad x(0) = x_0. \tag{3.18}$$

The spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}$ are called the error, state and signal spaces respectively. We define the set of trajectories $\mathcal{T}$ by

$$\mathcal{T} := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \begin{bmatrix} C(\mathbb{R}^+; \mathscr{X}) \\ L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) \end{bmatrix} : \exists\, x_0 \in \mathscr{X} \text{ such that (3.18) holds} \right\}. \tag{3.19}$$

The component x of a trajectory is understood as a solution of (3.18) as described in Remark 3.1.1. We define the set of trajectories from $x_0 \in \mathscr{X}$, $\mathcal{T}(x_0)$, by

$$\mathcal{T}(x_0) := \left\{ \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T} \;:\; x(0) = x_0 \right\}, \tag{3.20}$$

and define the set of externally generated trajectories $\mathcal{T}_{\text{ext}}$ by

$$\mathcal{T}_{\text{ext}} := \left\{ w \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{W}) : \exists\, x \in C(\mathbb{R}^+; \mathscr{X}) \text{ such that } \begin{bmatrix} x \\ w \end{bmatrix} \in \mathcal{T}(0) \right\}. \tag{3.21}$$

We define the set of stable externally generated trajectories $\mathcal{S}$ by

$$\mathcal{S} = \mathcal{T}_{\text{ext}} \cap L^2(\mathbb{R}^+; \mathscr{W}). \tag{3.22}$$

We call the pair consisting of the node (3.17) and the set of trajectories (3.19) an output-nulling system, which we denote by $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{on}}$.

Remark 3.1.9. The surjectivity of D in Definition 3.1.8 implies that for every $x_0 \in \mathscr{X}$ the corresponding set of trajectories from x_0, $\mathcal{T}(x_0)$, is non-empty. This can be proven directly, but it also follows as a consequence of Theorem 3.1.18.

3.1.2 Admissible decompositions

In this section we investigate when given a driving-variable or output-nulling system, it is possible to decompose the original signal space into an input space $\mathscr{U}$ and output space $\mathscr{Y}$ such that the signals of the driving-variable or output-nulling system are the trajectories of an input-state-output system.

Remark 3.1.10. Let $\mathscr{U}, \mathscr{Y}$ denote a direct sum decomposition of a finite-dimensional Hilbert space $\mathscr{W}$, which we denote by $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ and understand as $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ 0 \end{smallmatrix}\right] \oplus \left[\begin{smallmatrix} 0 \\ \mathscr{Y} \end{smallmatrix}\right]$. We identify $u \in \mathscr{U}$ with $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathscr{U} \\ 0 \end{smallmatrix}\right]$, etc. so that $w = u + y = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$. We let $\pi_{\mathscr{U}}^{\mathscr{Y}}$ ($\pi_{\mathscr{Y}}^{\mathscr{U}}$) denote the projection of $\mathscr{W}$ onto $\mathscr{U}$ ($\mathscr{Y}$) along $\mathscr{Y}$ ($\mathscr{U}$) and given an operator

$$T : \mathscr{X} \to \mathscr{W},$$

for some linear space $\mathscr{X}$ we write

$$T_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}} T, \quad T_{\mathscr{Y}} = \pi_{\mathscr{Y}}^{\mathscr{U}} T. \tag{3.23}$$

Definition 3.1.11. Given a driving-variable system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, let $\mathscr{U}$ and $\mathscr{Y}$ denote complementary subspaces of the signal space $\mathscr{W}$. We say that the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible for $\mathscr{W}$, or simply strongly admissible, if

$$(\pi_{\mathscr{U}}^{\mathscr{Y}} D)^{-1} : \mathscr{U} \to \mathscr{V}, \quad \text{exists.}$$

Given an output-nulling system with node $\left[\begin{smallmatrix} A' & B' \\ C' & D' \end{smallmatrix}\right]$ and a surjective operator $E \in B(\mathscr{W})$, we say that pair $\mathscr{U}$, $\mathscr{Y}$ of complementary subspaces is E-strongly admissible for $\mathscr{W}$, or simply E-strongly admissible, if

$$(D' E|_{\mathscr{Y}})^{-1} : \mathscr{V} \to \mathscr{Y}, \quad \text{exists.}$$

If $E = I$, the identity on $\mathscr{W}$, then we say that the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible instead of I-strongly admissible.

Remark 3.1.12.
1. In this chapter we have chosen to work mostly with driving-variable systems instead of output-nulling systems. We could have defined an E-strongly admissible pair for driving-variable systems, where E is now injective, but have no need for this. We need the notion of an E-strongly admissible pair for output-nulling systems as defined above for duality, which we address in Section 3.4. We comment that many of the following results are stated and proven from the point of view of driving-variable systems and the corresponding output-nulling versions have been omitted.

2. We include the word strongly in strongly admissible so as to contrast the concept with a weaker version, which we discuss in Section 3.6.3.

Lemma 3.1.13. *The space $\mathscr{U} := \operatorname{im} D$ and any complementary subspace is always strongly admissible for a driving-variable system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. Similarly, for an output-nulling system with node $\left[\begin{smallmatrix} A' & B' \\ C' & D' \end{smallmatrix}\right]$, the subspace of $\mathscr{W}$ denoted by $\mathscr{Y}$, that is naturally isomorphic to the quotient space ${}^{\mathscr{W}}\!/_{\ker D'}$, and any complementary subspace is always strongly admissible (that is, I-strongly admissible). In particular, for any given driving-variable or output-nulling system there is always at least one strongly admissible pair.*

Proof. The first claim follows by injectivity of D and the fact that any map surjects onto its image. The second claim follows from surjectivity of D' and the First Isomorphism Theorem. □

Remark 3.1.14. For driving-variable systems the space $\mathscr{U} := \operatorname{im} D$ is referred to in [4] as the canonical input-space.

Definition 3.1.15. Given a driving-variable system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and a strongly admissible pair $\mathscr{U}, \mathscr{Y}$ we define the derived $(\mathscr{U}, \mathscr{Y})$ input-state-output node by

$$\begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{3.24}$$

which we denote by $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$. Recall that $C_{\mathscr{U}}, C_{\mathscr{Y}}, D_{\mathscr{U}}$ and $D_{\mathscr{Y}}$ are as in (3.23). We call the corresponding input-state-output system the derived $(\mathscr{U}, \mathscr{Y})$ input-state-output system.

Definition 3.1.16. Given an output-nulling system with node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, a surjective operator $E \in B(\mathscr{W})$ and an E-strongly admissible pair $\mathscr{U}, \mathscr{Y}$, we define the E-derived input-state-output node by

$$\begin{bmatrix} A - (BE)|_{\mathscr{Y}}(DE)|_{\mathscr{Y}}^{-1}C & (BE)|_{\mathscr{U}} - (BE)|_{\mathscr{Y}}(DE)|_{\mathscr{Y}}^{-1}(DE)|_{\mathscr{U}} \\ (DE)|_{\mathscr{Y}}^{-1}C & (DE)|_{\mathscr{Y}}^{-1}(DE)|_{\mathscr{U}} \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix}, \tag{3.25}$$

which we denote by $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$. We call the corresponding input-state-output system the E-derived $(\mathscr{U}, \mathscr{Y})$ input-state-output system. If $E = I$, the identity on $\mathscr{W}$, then we call the I-derived $(\mathscr{U}, \mathscr{Y})$ system the derived $(\mathscr{U}, \mathscr{Y})$ system instead.

Remark 3.1.17.

1. The terms strongly admissible and derived system have two meanings, one for driving-variable systems and one for output-nulling systems. In what follows it will be made clear which meaning is being used (though it is also often clear from the context).

2. A driving-variable or output-nulling system may have many possible derived systems, but once we fix a strongly admissible pair $\mathscr{U}, \mathscr{Y}$ (and where appropriate

$E \in B(\mathscr{W})$), then the derived $(\mathscr{U}, \mathscr{Y})$ system is uniquely specified by its node as in Definition 3.1.15 or 3.1.16 respectively.

The following result is crucial in obtaining input-state-output systems from driving-variable and output-nulling systems as it states that the trajectories of a strongly derived input-state-output system are the same as those of the original system.

Theorem 3.1.18. *Given a driving-variable system with set of trajectories $\mathcal{T}_{\mathrm{dv}}$ and a strongly admissible pair $\mathscr{U}, \mathscr{Y}$, let $\mathcal{T}_{\mathrm{iso}}$ denote the set of trajectories of the derived $(\mathscr{U}, \mathscr{Y})$ input-state-output system. Then the map $T : \mathcal{T}_{\mathrm{iso}} \to \mathcal{T}_{\mathrm{dv}}$ given by*

$$T \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} = \begin{bmatrix} x \\ \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \end{bmatrix}, \tag{3.26}$$

is an isomorphism.

Given an output-nulling system with set of trajectories $\mathcal{T}_{\mathrm{on}}$, a surjective operator $E \in B(\mathscr{W})$ and an E-strongly admissible pair $\mathscr{U}, \mathscr{Y}$, let $\mathcal{T}'_{\mathrm{iso}}$ denote the set of trajectories of the E-derived $(\mathscr{U}, \mathscr{Y})$ input-state-output system. Then the map $T' : \mathcal{T}'_{\mathrm{iso}} \to \mathcal{T}_{\mathrm{on}}$ given by

$$T' \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} = \begin{bmatrix} x \\ E\left[\begin{smallmatrix} u \\ -y \end{smallmatrix}\right] \end{bmatrix}, \tag{3.27}$$

is an isomorphism.

Remark 3.1.19. Using the notation of Theorem 3.1.18, we see that $\mathcal{T}_{\mathrm{dv}}$ and $\mathcal{T}_{\mathrm{iso}}$ are isomorphic and that $\mathcal{T}_{\mathrm{on}}$ and $\mathcal{T}'_{\mathrm{iso}}$ are isomorphic. In what follows we will say that $\mathcal{T}_{\mathrm{dv}}$ and $\mathcal{T}_{\mathrm{iso}}$ are equal, understood in the sense of (3.26), i.e.

$$\begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}_{\mathrm{iso}} \iff \begin{bmatrix} x \\ \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \end{bmatrix} \in \mathcal{T}_{\mathrm{dv}}.$$

We adopt the corresponding convention for $\mathcal{T}_{\mathrm{on}}$ and $\mathcal{T}'_{\mathrm{iso}}$, so that

$$\begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}'_{\mathrm{iso}} \iff \begin{bmatrix} x \\ E\left[\begin{smallmatrix} u \\ -y \end{smallmatrix}\right] \end{bmatrix} \in \mathcal{T}_{\mathrm{on}}.$$

Proof of Theorem 3.1.18: We only prove the driving-variable case (3.26), as the output-nulling case is similar. Both results essentially follow by construction. Suppose the driving-variable system has node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. The direct sum decomposition $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ implies that every $w \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{W})$ can be written as $w = u + y = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$, where $u \in$

$L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{U})$ and $y \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{Y})$. As such, if x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\mathrm{dv}}$ then by definition there exists a $v \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{V})$ such that

$$\begin{aligned} \dot{x} &= Ax + Bv, \\ \begin{bmatrix} u \\ y \end{bmatrix} &= Cx + Dv = \begin{bmatrix} C_{\mathscr{U}} x \\ C_{\mathscr{Y}} x \end{bmatrix} + \begin{bmatrix} D_{\mathscr{U}} v \\ D_{\mathscr{Y}} v \end{bmatrix}. \end{aligned} \tag{3.28}$$

By definition of strongly admissible in the driving-variable case the operator $D_{\mathscr{U}}$ is invertible and hence we can eliminate v from (3.28) and obtain

$$\begin{aligned} \dot{x} &= (A - BD_{\mathscr{U}}^{-1} C_{\mathscr{U}})x + BD_{\mathscr{U}}^{-1} u, \\ y &= (C_{\mathscr{Y}} - D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} C_{\mathscr{U}})x + D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} u, \end{aligned} \tag{3.29}$$

so that x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\mathrm{iso}}$. Conversely, if x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\mathrm{iso}}$ then defining

$$v := D_{\mathscr{U}}^{-1} u - D_{\mathscr{U}}^{-1} C_{\mathscr{U}} x \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{V}),$$

and substituting back into (3.29) we recover (3.28). As such, x and $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}_{\mathrm{dv}}$, completing the proof. □

Corollary 3.1.20. *The set of stable externally generated trajectories of a driving-variable system with signal space $\mathscr{W}$ is a closed subspace of $L^2(\mathbb{R}^+;\mathscr{W})$.*

Proof. Let Σ and $\mathcal{S}$ denote the driving-variable system and its set of stable externally generated trajectories respectively. Let $\mathfrak{D}$ denote the input-output map of the derived $(\mathscr{U},\mathscr{Y})$ system of Σ, for some choice of strongly admissible pair $\mathscr{U}$, $\mathscr{Y}$ (which always exists by Lemma 3.1.13). A consequence of Theorem 3.1.18 is that

$$\mathcal{S} = \left\{ \begin{bmatrix} u \\ \mathfrak{D}u \end{bmatrix} \ : \ u \in L^2(\mathbb{R}^+;\mathscr{U}) \quad \text{such that} \quad \mathfrak{D}u \in L^2(\mathbb{R}^+;\mathscr{Y}) \right\}. \tag{3.30}$$

It is well-known from input-state-output theory that the set on the right hand side of (3.30) is closed and hence so is $\mathcal{S}$. □

It will sometimes be helpful later in this work to obtain a driving-variable system from an input-state-output system and we describe how we do so in the next lemma.

Lemma 3.1.21. *Every input-state-output system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ with input, state and output spaces $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ respectively gives rise to a driving-variable system with*

$\mathscr{V} = \mathscr{U}$ *and* $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$*, driving-variable node*

$$\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right] : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W} \end{bmatrix}, \tag{3.31}$$

and set of trajectories $\mathcal{T}$.

Proof. This is immediate from the definitions, noting that the operator $\left[\begin{smallmatrix} I \\ D \end{smallmatrix}\right] : \mathscr{V} \to \mathscr{W}$ is always injective. Note that the set of trajectories of (3.31) is really isomorphic to $\mathcal{T}$, but we view them as equal (see Remark 3.1.19). □

Lemma 3.1.22. *If* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ *is an input-state-output system, then the pair* $\mathscr{U}, \mathscr{Y}$ *is always a strongly admissible pair for the driving-variable system*

$$\Sigma = \left(\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right], \mathcal{T} \right)_{\text{dv}}$$

constructed in Lemma 3.1.21, and the derived $(\mathscr{U}, \mathscr{Y})$ *system of* Σ *is the input-state-output system* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$.

Proof. This is immediate from Definition 3.1.15, Theorem 3.1.18 and Lemma 3.1.21. □

The following lemma characterises strong admissibility of a direct sum decomposition of the signal space of a driving-variable system and is based on [4, Lemma 5.7].

Lemma 3.1.23. *Given a driving-variable system* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$*, let* $\mathscr{U}, \mathscr{Y}$ *denote a direct sum decomposition of the signal space* $\mathscr{W}$*. The following are equivalent.*

(i) The pair $\mathscr{U}, \mathscr{Y}$ *is strongly admissible.*

(ii) There exists a map $\tilde{D} : \mathscr{U} \to \mathscr{Y}$ *such that* $\operatorname{im} D$ *has the graph representation*

$$\operatorname{im} D = \mathcal{G}(\tilde{D}) = \left\{ \begin{bmatrix} u \\ \tilde{D}u \end{bmatrix} \; : \; u \in \mathscr{U} \right\}.$$

Given a strongly admissible pair $\mathscr{U}, \mathscr{Y}$*, any another direct sum decomposition* $\mathscr{U}_1, \mathscr{Y}_1$ *is strongly admissible if and only if* $\dim \mathscr{U} = \dim \mathscr{U}_1$.

Proof. *(i)* $\Rightarrow$ *(ii)*: Supposing that the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible, let D_{D} denote the feedthrough operator of the derived $(\mathscr{U}, \mathscr{Y})$ system. We prove that

$$\text{im } D = \mathcal{G}(D_{\text{D}}), \tag{3.32}$$

which is condition (ii) with $D_{\mathrm{D}} = \tilde{D}$. To see (3.32) note that by Theorem 3.1.18 the input-output pairs $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right])$ (with zero initial state) of the derived system are precisely the externally generated signals of the driving-variable system. By considering continuous inputs u (and hence continuous outputs y) at time $t = 0$ we see that

$$\begin{bmatrix} u(0) \\ D_{\mathrm{D}}u(0) \end{bmatrix} = \begin{bmatrix} u(0) \\ y(0) \end{bmatrix} = Dv(0), \tag{3.33}$$

for some continuous driving-variable v, and thus $\mathcal{G}(D_{\mathrm{D}}) \subseteq \operatorname{im} D$. By considering continuous driving-variables v and the corresponding continuous signal $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ we deduce from (3.33) that $\mathcal{G}(D_{\mathrm{D}}) \supseteq \operatorname{im} D$, which proves (3.32).

$(ii) \Rightarrow (i)$: Given $u \in \mathscr{U}$, by assumption (ii) there is a $v \in \mathscr{V}$ such that $Dv = \left[\begin{smallmatrix} u \\ \tilde{D}u \end{smallmatrix}\right]$ and so $D_{\mathscr{U}}v = u$. We infer that $D_{\mathscr{U}}$ is surjective. For injectivity, if $D_{\mathscr{U}}v = 0$ then by (ii), $Dv = \left[\begin{smallmatrix} 0 \\ \tilde{D}0 \end{smallmatrix}\right] = 0$ and as D is injective it follows that $v = 0$. Hence $D_{\mathscr{U}}$ is injective as well as surjective and thus invertible.

If $\mathscr{U}_1, \mathscr{Y}_1$ is strongly admissible then the map

$$(D_{\mathscr{U}_1})(D_{\mathscr{U}})^{-1} : \mathscr{U} \to \mathscr{U}_1$$

is a linear isomorphism, and so $\dim \mathscr{U} = \dim \mathscr{U}_1$.

Conversely, suppose $\dim \mathscr{U}_1 = \dim \mathscr{U}$. Then there is a linear isomorphism $F : \mathscr{W} \to \mathscr{W}$ such that $F|_{\mathscr{U}_1} : \mathscr{U}_1 \to \mathscr{U}$ and $F|_{\mathscr{Y}_1} : \mathscr{Y}_1 \to \mathscr{Y}$ are isomorphisms. As such it follows that

$$\begin{aligned}\left\{ \begin{bmatrix} u \\ D_{\mathrm{D}}u \end{bmatrix} : u \in \mathscr{U} \right\} = F^{-1} \left\{ \begin{bmatrix} u \\ D_{\mathrm{D}}u \end{bmatrix} : u \in \mathscr{U} \right\} &= \left\{ \begin{bmatrix} F^{-1}u \\ F^{-1}D_{\mathrm{D}}u \end{bmatrix} : u \in \mathscr{U} \right\} \\ &= \left\{ \begin{bmatrix} u_1 \\ F^{-1}D_{\mathrm{D}}Fu_1 \end{bmatrix} : u_1 \in \mathscr{U}_1 \right\}. \end{aligned} \tag{3.34}$$

Since the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible it follows from the proof of (ii) above that $\operatorname{im} D = \mathcal{G}(D_{\mathrm{D}})$, which when combined with (3.34) implies that

$$\operatorname{im} D = \left\{ \begin{bmatrix} u_1 \\ F^{-1}D_{\mathrm{D}}Fu_1 \end{bmatrix} : u_1 \in \mathscr{U}_1 \right\},$$

with $F^{-1}D_{\mathrm{D}}F : \mathscr{U}_1 \to \mathscr{Y}_1$. We conclude from (iii) that $\mathscr{U}_1, \mathscr{Y}_1$ is a strongly admissible pair. □

Remark 3.1.24. In Willems & Trentelman [102, p. 55] the input cardinality $\mathtt{m}(\mathfrak{B})$ of a behavior $\mathfrak{B}$ is defined as the maximal number of unconstrained components of $w \in \mathfrak{B}$. It is stated [102, p. 60] that the input cardinality $\mathtt{m}(\mathfrak{B})$ is equal to the dimension of the input space of any input-state-output representation of $\mathfrak{B}$. Similarly, the output cardinality is defined as $\mathtt{w} - \mathtt{m}(\mathfrak{B})$, where $w \in \mathfrak{B}$ has $\mathtt{w}$ components. Since all linear differential behaviours $\mathfrak{B}$ admit a driving-variable representation, Lemma 3.1.13 and Lemma 3.1.23 show that the input cardinality of the behavior described by a driving-variable system is $\dim(\operatorname{im} D)$. Moreover, by Lemma 3.1.23 if $\mathscr{U}, \mathscr{Y}$ is a strongly admissible pair then necessarily $\dim \mathscr{U}$ equals the input cardinality.

3.1.3 Minimality

Definition 3.1.25. A system (input-state-output, driving-variable or output-nulling) with state space $\mathscr{X}$ and set of trajectories $\mathcal{T}$, is said to be minimal if supposing $\mathcal{T} = \mathcal{T}'$ for another such system with state space $\mathscr{X}'$ it follows that $\dim \mathscr{X} \leq \dim \mathscr{X}'$.

Remark 3.1.26. The above definition in the input-state-output case is consistent with the usual definition.

Lemma 3.1.27. *A driving-variable system is minimal if and only if for every strongly admissible pair $\mathscr{U}, \mathscr{Y}$ the derived $(\mathscr{U}, \mathscr{Y})$ system is minimal.*

Proof. Assume that the driving-variable system $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ with state space $\mathscr{X}$ is minimal and let $\mathscr{U}, \mathscr{Y}$ denote a strongly admissible pair. We seek to prove that the derived $(\mathscr{U}, \mathscr{Y})$ system, denoted $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$, is minimal. Suppose the input-state-output system $(\left[\begin{smallmatrix} A'_{\mathrm{D}} & B'_{\mathrm{D}} \\ C'_{\mathrm{D}} & D'_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ has state space $\mathscr{X}'$. By Lemma 3.1.22, $(\left[\begin{smallmatrix} A'_{\mathrm{D}} & B'_{\mathrm{D}} \\ C'_{\mathrm{D}} & D'_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ is the derived $(\mathscr{U}, \mathscr{Y})$ system of the driving-variable system

$$\left(\left[\begin{array}{c|c} A'_{\mathrm{D}} & B'_{\mathrm{D}} \\ \hline 0 & I \\ C'_{\mathrm{D}} & D'_{\mathrm{D}} \end{array} \right], \mathcal{T} \right)_{\mathrm{dv}},$$

which also has state space $\mathscr{X}'$. Minimality of Σ implies that

$$\dim \mathscr{X} \leq \dim \mathscr{X}', \tag{3.35}$$

and hence the derived $(\mathscr{U}, \mathscr{Y})$ input-state-output system is minimal.

Conversely, given the driving-variable system $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ with state space $\mathscr{X}$, assume that the derived $(\mathscr{U}, \mathscr{Y})$ system for any strongly admissible pair $\mathscr{U}, \mathscr{Y}$ is minimal. Suppose the driving-variable node $\left[\begin{smallmatrix} A' & B' \\ C' & D' \end{smallmatrix}\right]$ with state space $\mathscr{X}'$ has the same set of trajectories $\mathcal{T}$ as Σ. Denote this system by Σ'. We seek to prove that (3.35) holds. The result will follow if we can establish that $\mathscr{U}, \mathscr{Y}$ is a strongly admissible pair for Σ'.

Since Σ and Σ' share the same set of trajectories, they therefore also have the same set of externally generated trajectories $\mathcal{T}_{\mathrm{ext}}$. By considering (continuous) signals belonging to $\mathcal{T}_{\mathrm{ext}}$ at time $t = 0$ it follows that $\operatorname{im} D' = \operatorname{im} D$, and as $\mathscr{U}, \mathscr{Y}$ is an admissible pair for Σ it follows from Lemma 3.1.23 (*ii*) that

$$\operatorname{im} D' = \operatorname{im} D = \left\{ \begin{bmatrix} u \\ \tilde{D}u \end{bmatrix} \,:\, u \in \mathscr{U} \right\}. \tag{3.36}$$

for some operator $\tilde{D} : \mathscr{U} \to \mathscr{Y}$. In light of (3.36) and Lemma 3.1.23 (*ii*) we see that $\mathscr{U}, \mathscr{Y}$ is a strongly admissible pair for Σ', which completes the proof. □

3.1.4 The gap metric

We recall the definition of the gap metric, see also Kato [43, p.197].

Definition 3.1.28. For $\mathscr{M}, \mathscr{N}$ non-empty closed subspaces of a Banach space $\mathscr{X}$, the directed gap δ between $\mathscr{M}$ and $\mathscr{N}$ is given by

$$\delta(\mathscr{M}, \mathscr{N}) = \sup_{\substack{m \in \mathscr{M} \\ \|m\|=1}} \operatorname{dist}(m, \mathscr{N}),$$

with the conventions

$$\delta(\{0\}, \mathscr{N}) = 0, \quad \delta(\mathscr{M}, \{0\}) = 1, \quad \text{for } \mathscr{M} \neq \{0\}.$$

The gap $\hat{\delta}$ between $\mathscr{M}$ and $\mathscr{N}$ is then defined as

$$\hat{\delta}(\mathscr{M}, \mathscr{N}) = \max\{\delta(\mathscr{M}, \mathscr{N}), \delta(\mathscr{N}, \mathscr{M})\}.$$

In general $\hat{\delta}$ is not a metric on the set of closed subspaces of a Banach space, as it does not satisfy the triangle inequality. However, when $\mathscr{X}$ is a Hilbert space it can be shown that

$$\hat{\delta}(\mathscr{M}, \mathscr{N}) = ||P_{\mathscr{M}} - P_{\mathscr{N}}||, \tag{3.37}$$

where $P_{\mathscr{M}}, P_{\mathscr{N}}$ are the orthogonal projections of $\mathscr{X}$ onto $\mathscr{M}$ and $\mathscr{N}$ respectively. In this instance $\hat{\delta}$ is a metric. We now define the gap between two driving-variable systems.

Definition 3.1.29. Let Σ_1 and Σ_2 denote two driving-variable systems with the same signal space and sets of stable externally generated trajectories $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. We define the gap between Σ_1 and Σ_2 as

$$\hat{\delta}(\Sigma_1, \Sigma_2) := \hat{\delta}(\mathcal{S}_1, \mathcal{S}_2).$$

Remark 3.1.30. The gap between two driving-variable systems is well-defined as the sets of stable externally generated trajectories are closed subspaces of $L^2(\mathbb{R}^+; \mathscr{W})$ by Corollary 3.1.20.

We shall also need the definition of the gap between two closed operators, as in [43, p. 201].

Definition 3.1.31. For $\mathscr{X}_i$ Hilbert spaces and closed linear operators $S, T : \mathscr{X}_1 \to \mathscr{X}_2$, the gap between S and T is defined as

$$\hat{\delta}(S, T) := \hat{\delta}(\mathcal{G}(S), \mathcal{G}(T)). \tag{3.38}$$

The next result is taken from [43] and is an important bound that we shall make use of later.

Theorem 3.1.32. *For $\mathscr{X}_i$ Hilbert spaces and bounded linear operators $S, T : \mathscr{X}_1 \to \mathscr{X}_2$ we have*

$$\hat{\delta}(S, T) \leq \|S - T\|. \tag{3.39}$$

Proof. See [43, Theorem 2.14]. $\square$

3.2 Finite-dimensional indefinite inner-product spaces

In order to describe what we mean by a dissipative system we shall need the concept of an indefinite inner-product. In this section we collect some elementary definitions and results on complex finite-dimensional indefinite inner-product spaces. Three supplementary references for this material are Bognár [10] and Gohberg *et al.* [30], [31].

Definition 3.2.1. Let $\mathscr{W}$ denote a finite-dimensional linear space. A function $[\cdot, \cdot] : \mathscr{W} \times \mathscr{W} \to \mathbb{C}$ is called an indefinite (non-degenerate) inner-product on $\mathscr{W}$ if the following axioms are satisfied:

(1) Linearity in the second argument

$$[x, \alpha y_1 + \beta y_2] = \alpha[x, y_1] + \beta[x, y_2]$$

for all $x, y_1, y_2 \in \mathscr{W}$ and all $\alpha, \beta \in \mathbb{C}$;

(2) Antisymmetry

$$[x, y] = \overline{[y, x]};$$

for all $x, y \in \mathscr{W}$,

(3) Nondegeneracy; if $[x, y] = 0$ for all $y \in \mathscr{W}$, then $x = 0$.

Note that in contrast to a definite inner-product, $[x, x] < 0$ can occur. We call $\mathscr{W}$ equipped with $[\cdot, \cdot]$ an indefinite inner-product space, and we denote it by $(\mathscr{W}, [\cdot, \cdot])$ or sometimes just $\mathscr{W}$.

Lemma 3.2.2. *If $(\mathscr{W}, [\cdot,\cdot])$ is a finite-dimensional indefinite inner-product space then there exists a definite inner-product $\langle\cdot,\cdot\rangle$ on $\mathscr{W}$ and a unique unitary self-adjoint operator E (with respect to $\langle\cdot,\cdot\rangle$) such that*

$$[x, y] = \langle Ex, y\rangle, \quad \forall\, x, y \in \mathscr{W}. \tag{3.40}$$

We say that $\langle\cdot,\cdot\rangle$ is the definite *inner-product and E is the signature operator induced by the* indefinite *inner-product $[\cdot,\cdot]$.*

Proof. The proof extends the arguments of [31, p. 8]. Let $(\cdot,\cdot) : \mathscr{W} \times \mathscr{W} \to \mathbb{C}$ denote some (definite) inner-product on $\mathscr{W}$, which is always possible to find as $\mathscr{W}$ is finite-dimensional. For fixed $x \in \mathscr{W}$ consider the linear functional

$$\mathscr{W} \ni y \mapsto [x, y],$$

which is certainly bounded. By the Riesz Representation Theorem there exists a unique $z_x \in \mathscr{W}$ such that

$$[x, y] = (z_x, y), \quad \forall y \in \mathscr{W}. \tag{3.41}$$

Define $H : \mathscr{W} \to \mathscr{W}$ by $Hx = z_x$, where x and z_x are as in (3.41). Since for each $x \in \mathscr{W}$, z_x is uniquely determined it follows that H is well-defined and in fact injective. It is easy to see that H is linear and hence (as $\mathscr{W}$ is finite-dimensional) H is an isomorphism. For $x, y \in \mathscr{W}$ we have

$$(Hx, y) = [x, y] = \overline{[y, x]} = \overline{(Hy, x)} = (x, Hy),$$

and so H is self-adjoint with respect to $(\cdot,\cdot)$. Let $H = |H|E$ denote the polar decomposition of H, where $|H|$ is positive and self-adjoint and E is unitary, both with respect to $(\cdot,\cdot)$. For this proof let the superscript $*$ denote the adjoint with respect to $(\cdot,\cdot)$. Define the new inner-product $\langle\cdot,\cdot\rangle : \mathscr{W} \times \mathscr{W} \to \mathbb{C}$ by

$$\langle x, y\rangle := (|H|x, y), \quad \forall x, y \in \mathscr{W}. \tag{3.42}$$

We see from (3.41) and (3.42) that for $x, y \in \mathscr{W}$

$$[x, y] = (Hx, y) = (|H|Ex, y) = \langle Ex, y\rangle,$$

which is (3.40). It remains to see that E is self-adjoint and unitary with respect to $\langle\cdot,\cdot\rangle$. Since $H = H^*$ in Kato [43, p. 335] it is proven that $E = E^*$ and thus

$$HE = HE^* = |H|EE^* = |H|. \tag{3.43}$$

Therefore, for $x, y \in \mathscr{W}$

$$\langle Ex, Ey \rangle = (Ex, |H|Ey) = (Ex, Hy) = (HEx, y) = (|H|x, y) = \langle x, y \rangle, \tag{3.44}$$

where we have used (3.43). Equality (3.44) proves that E is unitary. For self-adjointness, let $x, y \in \mathscr{W}$

$$\langle x, Ey \rangle = (x, |H|Ey) = (x, Hy) = (Hx, y) = (|H|Ex, y) = \langle Ex, y \rangle, \tag{3.45}$$

where we have used that $H = H^*$ and $|H| = |H|^*$. Equality (3.45) proves that E is self-adjoint with respect to $\langle \cdot, \cdot \rangle$. Finally, E is unique from the uniqueness of H and of the polar decomposition. $\square$

Remark 3.2.3. In Section 3.3 we shall consider driving-variable systems where $\mathscr{W}$ is an indefinite inner-product space. We remark that in this instance we use the *definite* inner-product $\langle \cdot, \cdot \rangle$ (induced by the indefinite inner-product $[\cdot, \cdot]$) in the theory of driving-variable systems established in Section 3.1. In particular, $L^2(\mathbb{R}^+; \mathscr{W})$ is a Hilbert space when $\mathscr{W}$ is equipped with $\langle \cdot, \cdot \rangle$.

Definition 3.2.4. A subspace $\mathscr{S}$ of an indefinite inner-product space $(\mathscr{W}, [\cdot, \cdot])$ is called non-negative (respectively, neutral, non-positive) if $[x, x] \geq 0$ ($[x, x] = 0, [x, x] \leq 0$) for all $x \in \mathscr{S}$. A non-negative subspace is called maximal if it is not the proper subset of another non-negative subspace, with similar definitions for maximal neutral and non-positive subspaces.

Definition 3.2.5. Given an indefinite inner-product space $(\mathscr{W}, [\cdot, \cdot])$ we say that $\mathscr{W}_+, \mathscr{W}_- \subseteq \mathscr{W}$ is a fundamental decomposition of $\mathscr{W}$, denoted $\mathscr{W} = \mathscr{W}_+ [+] - \mathscr{W}_-$ if,

(1) $\mathscr{W}_+$ equipped with $[\cdot, \cdot]_{|\mathscr{W}_+}$ and $\mathscr{W}_-$ equipped with $-[\cdot, \cdot]_{|\mathscr{W}_-}$ are Hilbert spaces.

(2) $\mathscr{W}$ is a direct sum decomposition of $\mathscr{W}_+$ and $\mathscr{W}_-$, orthogonal with respect to $[\cdot, \cdot]$.

Note that if $\mathscr{W} = \mathscr{W}_+ [+] - \mathscr{W}_-$ then by (1) $\mathscr{W}_+$ is non-negative and $\mathscr{W}_-$ is non-positive. Fundamental decompositions are in general not unique.

Definition 3.2.6. For a subspace $\mathscr{S}$ of an indefinite inner-product space $(\mathscr{W}, [\cdot, \cdot])$ we denote by $\mathscr{S}^{[\perp]}$ the orthogonal companion with respect to $[\cdot, \cdot]$, which is defined as

$$\mathscr{S}^{[\perp]} = \{w \in \mathscr{W} \ : \ [w, v] = 0, \ \forall \, v \in \mathscr{S}\}.$$

Note that $\mathscr{S} \cap \mathscr{S}^{[\perp]} \neq \{0\}$ in general.

Lemma 3.2.7. *If $\mathscr{W}_+$, $\mathscr{W}_-$ is a fundamental decomposition of an indefinite inner-product space $(\mathscr{W}, [\cdot, \cdot])$ with signature operator E, then* *(i)* $\mathscr{S} \subseteq \mathscr{W}$ *is non-negative if and only if* $\mathscr{S} = \mathcal{G}(T)$*, for* $T : \mathscr{W}_+ \supseteq D(T) \to \mathscr{W}_-$ *a linear contraction, where* $D(T)$ *is the domain of* T*. Additionally* $\mathscr{S}$ *is maximal non-negative if and only if* $D(T) = \mathscr{W}_+$*.*

(ii) $\mathscr{S} \subseteq \mathscr{W}$ *is maximal non-negative if and only if* $\mathscr{S}$ *is non-negative and* $\mathscr{S}^{[\perp]}$ *is non-positive.*

(iii) The dimension of any maximal non-negative (non-positive) subspace is equal to the multiplicity of 1 (-1) *as an eigenvalue of* E*. Hence any two maximal non-negative (non-positive) subspaces are isomorphic.*

(iv) The dimensions of the non-negative parts of any two fundamental decompositions are the same.

Proof. For parts (i) and (ii) see [10], namely Theorem 4.2, Theorem 4.4 and Lemma 4.5 on p. 105-106. For part (iii) see [30, Theorem 1.3, p. 15]. Part (iv) follows immediately from (iii). $\square$

Corollary 3.2.8. *For a driving-variable system with node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$*, indefinite inner-product signal space* $\mathscr{W}$ *and signature operator* E*, a fundamental decomposition* $\mathscr{W}_+$*,* $\mathscr{W}_-$ *of* $\mathscr{W}$ *is a strongly admissible pair for* $\mathscr{W}$ *if and only if*

$$\dim \mathscr{W}_+ = \sigma_+(E) = \dim(\operatorname{im} D).$$

Proof. By Lemma 3.1.23, the direct sum decomposition $\mathscr{W}_+$, $\mathscr{W}_-$ of $\mathscr{W}$ is strongly admissible if and only if $\dim \mathscr{W}_+ = \dim(\operatorname{im} D)$ as by Lemma 3.1.13 the pair $\operatorname{im} D$ and any complementary subspace is always strongly admissible. That $\dim \mathscr{W}_+ = \sigma_+(E)$ follows from Lemma 3.2.7 (iii) above. $\square$

3.3 Dissipative systems

We are now in position to define dissipativity for the systems we considered in Section 3.1.1.

Definition 3.3.1. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ denote an input-state-output system and let

$$[\cdot,\cdot] : \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix} \to \mathbb{C},$$

denote an indefinite inner-product on $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$. We say that $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ is state-signal dissipative with respect to $[\cdot,\cdot]$ if there exists a positive, self-adjoint operator P on $\mathscr{X}$

such that

$$\int_0^t \left[\left[\begin{smallmatrix} u(s) \\ y(s) \end{smallmatrix}\right], \left[\begin{smallmatrix} u(s) \\ y(s) \end{smallmatrix}\right]\right] \, ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px(0), x(0)\rangle_{\mathscr{X}},$$
$$\forall\, t \geq 0,\ \forall \begin{bmatrix} x \\ \hline u \\ y \end{bmatrix} \in \mathcal{T}. \tag{3.46}$$

We call $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ signal dissipative with respect to $[\cdot,\cdot]$ if

$$\int_0^t \left[\left[\begin{smallmatrix} u(s) \\ y(s) \end{smallmatrix}\right], \left[\begin{smallmatrix} u(s) \\ y(s) \end{smallmatrix}\right]\right] \, ds \geq 0, \quad \forall\, t \geq 0, \forall \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathcal{T}_{\text{ext}}. \tag{3.47}$$

Remark 3.3.2. The above definition is similar in effect to equipping an input-state-output system with a supply rate in the language of [99].

Definition 3.3.3. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$ denote a driving-variable system with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$. We say that $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$ is state-signal dissipative if there exists a positive, self-adjoint operator P on $\mathscr{X}$ such that

$$\int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px(0), x(0)\rangle_{\mathscr{X}},$$
$$\forall\, t \geq 0,\ \forall \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}. \tag{3.48}$$

We call $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$ signal dissipative if

$$\int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds \geq 0, \quad \forall\, t \geq 0, \forall\, w \in \mathcal{T}_{\text{ext}}. \tag{3.49}$$

Definition 3.3.4. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{on}}$ denote an output-nulling system with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$. We say that $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{on}}$ is state-signal dissipative if there exists a positive, self-adjoint operator P on $\mathscr{X}$ such that

$$\int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds \geq \langle Px(t), x(t)\rangle_{\mathscr{X}} - \langle Px(0), x(0)\rangle_{\mathscr{X}},$$
$$\forall\, t \geq 0,\ \forall \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}. \tag{3.50}$$

We call $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{on}}$ signal dissipative if

$$\int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds \geq 0, \quad \forall\, t \geq 0, \forall\, w \in \mathcal{T}_{\text{ext}}. \tag{3.51}$$

Remark 3.3.5. Observe that for all three types of systems (input-state-output, driving-variable and output-nulling) the state x is continuous and so the point evaluations in (3.46), (3.48) and (3.50) make sense. In the input-state-output case let E' denote the signature operator of the indefinite inner-product on $\left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right]$ and let x and $\left[\begin{smallmatrix}u\\ y\end{smallmatrix}\right]$ denote the components of a trajectory. For $t \geq 0$ we have

$$\left|\int_0^t \left[\left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right], \left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right]\right] \, ds\right| \leq \int_0^t \left|\left\langle \left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right], E' \left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right]\right\rangle\right| \, ds,$$

which by the Hölder inequality gives,

$$\begin{aligned}\left|\int_0^t \left[\left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right], \left[\begin{smallmatrix}u(s)\\ y(s)\end{smallmatrix}\right]\right] \, ds\right| &\leq \| \left[\begin{smallmatrix}u\\ y\end{smallmatrix}\right] \|_{L^2(0,t)} \cdot \|E' \left[\begin{smallmatrix}u\\ y\end{smallmatrix}\right] \|_{L^2(0,t)} \\ &\leq \| \left[\begin{smallmatrix}u\\ y\end{smallmatrix}\right] \|^2_{L^2(0,t)} < \infty, \quad \text{as } \|E'\| = 1.\end{aligned}$$

Similarly, in both the driving-variable and output nulling cases, if $\left[\begin{smallmatrix}x\\ w\end{smallmatrix}\right] \in \mathcal{T}$ then for $t \geq 0$ by the Hölder inequality we have

$$\left|\int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds\right| \leq \int_0^t |\langle w(s), Ew(s)\rangle| \, ds \leq \|w\|^2_{L^2(0,t)} < \infty,$$

where E is the signature operator of the indefinite inner-product $[\cdot,\cdot]_{\mathscr{W}}$ on $\mathscr{W}$. We conclude that dissipativity is well-defined.

Remark 3.3.6. From the definitions of dissipativity we see immediately that in all three cases signal dissipativity is a necessary condition for state-signal dissipativity. In Theorem 3.5.5 the converse implication is addressed.

The next three results describe the relationships between dissipativity of driving-variable systems and their derived input-state-output systems.

Proposition 3.3.7. *If the input-state-output system $(\left[\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right], \mathcal{T})_{\text{iso}}$ is state-signal or signal dissipative then the driving-variable system*

$$\left(\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right], \mathcal{T}\right)_{\text{dv}},$$

constructed in Proposition 3.1.21 is dissipative in the same sense, where the indefinite inner-product on $\mathscr{W}$ is that put on $\left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right]$.

Proof. This is immediate from the definitions. $\square$

Proposition 3.3.8. *Let $(\left[\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$ denote a driving-variable system with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ and assume that the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible. If $(\left[\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$ is state-signal or signal dissipative then the derived $(\mathscr{U}, \mathscr{Y})$ system is dissipative in the same sense, with respect to the indefinite inner-product $[\cdot,\cdot]_{\mathscr{W}}$.*

Proof. This again follows from the definitions and the fact that by Theorem 3.1.18 the original driving-variable system and the derived $(\mathscr{U}, \mathscr{Y})$ system have the same trajectories. $\square$

Theorem 3.3.9. *Let* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *denote a driving-variable system, with indefinite inner-product signal space* $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ *and let* $\mathscr{U}_{\mathrm{br}}, \mathscr{Y}_{\mathrm{br}}$ *denote a fundamental decomposition. If* $\mathscr{U}_{\mathrm{br}}, \mathscr{Y}_{\mathrm{br}}$ *is strongly admissible and* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *is signal dissipative then the derived* $(\mathscr{U}_{\mathrm{br}}, \mathscr{Y}_{\mathrm{br}})$ *system is bounded real.*

If additionally $\sigma_+(E) = \sigma_-(E)$ *then there exists a strongly admissible pair* $\mathscr{U}_{\mathrm{pr}}, \mathscr{Y}_{\mathrm{pr}}$ *such that the derived* $(\mathscr{U}_{\mathrm{pr}}, \mathscr{Y}_{\mathrm{pr}})$ *system is positive real.*

Proof. The first claim is an immediate consequence of Proposition 3.3.8 and the Bounded Real Lemma. In more detail, since $\mathscr{U}_{\mathrm{br}}$, $\mathscr{Y}_{\mathrm{br}}$ is a fundamental decomposition of the signal space $\mathscr{W}$ it follows that for any $w \in \mathscr{W}$, there exists $u \in \mathscr{U}_{\mathrm{br}}$ and $y \in \mathscr{Y}_{\mathrm{br}}$ such that $w = u + y = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ and moreover

$$\begin{aligned} [w,w]_{\mathscr{W}} &= [u+y, u+y]_{\mathscr{W}} = [u,u]_{\mathscr{W}} + [y,y]_{\mathscr{W}}, \qquad \text{by orthogonality} \\ &= \langle u,u\rangle_{\mathscr{U}_{\mathrm{br}}} - \langle y,y\rangle_{\mathscr{Y}_{\mathrm{br}}}, \end{aligned} \tag{3.52}$$

where $\langle\cdot,\cdot\rangle_{\mathscr{U}_{\mathrm{br}}}$ and $\langle\cdot,\cdot\rangle_{\mathscr{Y}_{\mathrm{br}}}$ denote the definite inner-products on $\mathscr{U}_{\mathrm{br}}$ and $\mathscr{Y}_{\mathrm{br}}$ induced by $[\cdot,\cdot]_{\mathscr{W}}$ respectively. Therefore, by Proposition 3.3.8 and (3.52), for all externally generated trajectories $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ of the derived $(\mathscr{U}_{\mathrm{br}}, \mathscr{Y}_{\mathrm{br}})$ system we have

$$\int_0^t \|u(s)\|^2 - \|y(s)\|^2 \; ds \geq 0, \quad \forall\, t \geq 0.$$

From the Bounded Real Lemma we infer that the derived $(\mathscr{U}_{\mathrm{br}}, \mathscr{Y}_{\mathrm{br}})$ system is bounded real.

With respect to the decomposition $\mathscr{W} = \mathscr{U}_{\mathrm{br}}[+] - \mathscr{Y}_{\mathrm{br}}$ the signature operator E has the block diagonal form

$$\begin{bmatrix} I_{\sigma_+(E)} & 0 \\ 0 & -I_{\sigma_-(E)} \end{bmatrix},$$

If we additionally assume that $\sigma_+(E) = \sigma_-(E)$ then under the transformation

$$\begin{bmatrix} \mathscr{U}_{\mathrm{pr}} \\ \mathscr{Y}_{\mathrm{pr}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & -I \\ I & I \end{bmatrix} \begin{bmatrix} \mathscr{U}_{\mathrm{br}} \\ \mathscr{Y}_{\mathrm{br}} \end{bmatrix},$$

$\mathscr{W}$ is the direct sum of $\mathscr{U}_{\mathrm{pr}}$ and $\mathscr{Y}_{\mathrm{pr}}$ and the signature operator E with respect to this decomposition has the block form

$$\begin{bmatrix} 0 & I_{\sigma_+(E)} \\ I_{\sigma_+(E)} & 0 \end{bmatrix},$$

and $\dim \mathscr{U}_{\rm pr} = \sigma_+(E) = \dim \mathscr{U}_{\rm br}$. As $\mathscr{U}_{\rm br}$ and $\mathscr{Y}_{\rm br}$ is strongly admissible, Lemma 3.1.23 implies that $\mathscr{U}_{\rm pr}$, $\mathscr{Y}_{\rm pr}$ is a strongly admissible pair. Moreover, with respect to this decomposition of $\mathscr{W}$ it follows that for $w \in \mathscr{W}$ there exists $u^\times \in \mathscr{U}_{\rm pr}$, $y^\times \in \mathscr{Y}_{\rm pr}$ such that $w = u^\times + y^\times = \left[\begin{smallmatrix} u^\times \\ y^\times \end{smallmatrix}\right]$ and we obtain the impedance supply rate

$$[w, w]_{\mathscr{W}} = [u^\times + y^\times, u^\times + y^\times]_{\mathscr{W}} = 2 \operatorname{Re} \langle u^\times, y^\times \rangle_{\mathscr{W}}. \tag{3.53}$$

That the derived $(\mathscr{U}_{\rm pr}, \mathscr{Y}_{\rm pr})$ system is positive real now follows from Proposition 3.3.8, equality (3.53) and the Positive Real Lemma. $\square$

3.4 Dual systems

Here we describe the duals of the state-space systems introduced in Section 3.1. We start with the familiar input-state-output case.

Definition 3.4.1. Given an input-state-output node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ with input, state and output spaces $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ respectively, we define the dual input-state-output node as the operator

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix}. \tag{3.54}$$

We denote by $\mathcal{T}^*$ the set of trajectories as in Definition 3.1.2, only now corresponding to the input-state-output node $\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$, and call $(\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right], \mathcal{T}^*)_{\rm iso}$ the dual input-state-output system.

The dual of a driving-variable node is an output-nulling node and vice versa. In order to formulate these definition we need some more notation.

Definition 3.4.2. Let $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ denote an indefinite inner-product space. We define $(\mathscr{W}^*, [\cdot,\cdot]_{\mathscr{W}^*})$ as the linear space $\mathscr{W}$ equipped with the indefinite inner-product $-[\cdot,\cdot]_{\mathscr{W}}$, and call $\mathscr{W}^*$ the anti-space of $\mathscr{W}$. Moreover, given Hilbert spaces $\mathscr{X}$ and $\mathscr{V}$ and operators

$$C : \mathscr{X} \to \mathscr{W}, \quad D : \mathscr{V} \to \mathscr{W},$$

we define

$$C^\dagger : \mathscr{W}^* \to \mathscr{X}, \quad D^\dagger : \mathscr{W}^* \to \mathscr{V},$$

as the adjoint maps, taken with respect to the Hilbert space inner-products on $\mathscr{X}$ (or $\mathscr{V}$) and the indefinite inner-product on $\mathscr{W}^*$. Thus $C^\dagger$ and $D^\dagger$ are such that

$$[w, Cx]_{\mathscr{W}^*} = \langle C^\dagger w, x\rangle_{\mathscr{X}}, \quad \forall\, x \in \mathscr{X},\ \forall\, w \in \mathscr{W}, \tag{3.55}$$

$$[w, Dv]_{\mathscr{W}^*} = \langle D^\dagger w, v\rangle_{\mathscr{V}}, \quad \forall\, v \in \mathscr{V},\ \forall\, w \in \mathscr{W}. \tag{3.56}$$

Definition 3.4.3. Given a driving-variable node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$, we define the dual output-nulling node by

$$\begin{bmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{W}^* \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix}, \tag{3.57}$$

which has error, state and signal spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}^*$ respectively. We denote by $\mathcal{T}^*$ the set of trajectories as in Definition 3.1.8, only now corresponding to the output-nulling node $\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right]$, and call $(\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\text{on}}$ the dual output-nulling system.

Definition 3.4.4. Given an output-nulling node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$, we define the dual driving-variable node by

$$\begin{bmatrix} A^* & C^\dagger \\ B^* & D^\dagger \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{V} \end{bmatrix} \to \begin{bmatrix} \mathscr{X} \\ \mathscr{W}^* \end{bmatrix}, \tag{3.58}$$

with driving-variable, state and signal spaces $\mathscr{V}$, $\mathscr{X}$ and $\mathscr{W}^*$ respectively. We denote by $\mathcal{T}^*$ the set of trajectories as in Definition 3.1.6, only now corresponding to the driving-variable node $\left[\begin{smallmatrix} A^* & C^\dagger \\ B^* & D^\dagger \end{smallmatrix}\right]$, and call $(\left[\begin{smallmatrix} A^* & C^\dagger \\ B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\text{dv}}$ the dual driving-variable system.

Proposition 3.4.5. *Given a driving-variable (output-nulling) system* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$*, let* $\mathcal{T}^*$ *denote the set of externally generated trajectories of the dual output-nulling (driving-variable) system. The following orthogonality relation holds*

$$\int_0^t [w(s), w_*(t-s)]_{\mathscr{W}}\, ds = 0, \quad \forall\, w \in \mathcal{T}_{\text{ext}},\ \forall\, w_* \in \mathcal{T}^*_{\text{ext}},\ \forall\, t \geq 0.$$

Proof. As both proofs are similar, we prove the case when $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})$ is a driving-variable system and $\mathcal{T}^*$ denotes the set of trajectories of the dual output-nulling system. For $w \in \mathcal{T}_{\text{ext}}$, $w_* \in \mathcal{T}^*_{\text{ext}}$ and $t \geq 0$ a calculation shows that

$$\begin{aligned}
\int_0^t [w(s), w_*(t-s)]_{\mathscr{W}}\, ds &= \int_0^t [Cx(s) + Dv(s), w_*(t-s)]_{\mathscr{W}}\, ds \\
&= \int_0^t -\langle x(s), C^\dagger w_*(t-s)\rangle_{\mathscr{X}} - \langle v(s), D^\dagger w_*(t-s)\rangle_{\mathscr{V}}\, ds \\
&= \int_0^t \langle x(s), \dot{x}_*(t-s) - A^* x_*(t-s)\rangle_{\mathscr{X}} - \langle v(s), B^* x_*(t-s)\rangle_{\mathscr{V}}\, ds \\
&= \int_0^t \langle x(s), \dot{x}_*(t-s)\rangle_{\mathscr{X}} - \langle Ax(s) + Bv(s), x_*(t-s)\rangle_{\mathscr{X}}\, ds \\
&= -\int_0^t \frac{d}{ds}\langle x(s), x_*(t-s)\rangle_{\mathscr{X}}\, ds = \langle x(0), x_*(t)\rangle_{\mathscr{X}} - \langle x(t), x_*(0)\rangle_{\mathscr{X}} \\
&= 0.
\end{aligned}$$

□

Lemma 3.4.6. *Let $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ denote an indefinite inner-product space with signature operator E and let $\mathscr{X}$ and $\mathscr{V}$ denote Hilbert spaces. For operators*

$$C : \mathscr{X} \to \mathscr{W}, \quad D : \mathscr{V} \to \mathscr{W},$$

the Hilbert space adjoints and indefinite inner-product space adjoints are related by

$$D^\dagger = -D^* E, \qquad C^\dagger = -C^* E, \tag{3.59}$$

where $C^\dagger, D^\dagger$ are as in Definition 3.4.2. Furthermore, if $\mathscr{U}$ and $\mathscr{Y}$ are a direct sum decomposition for $\mathscr{W}$ then

$$\begin{aligned} C^*|_{\mathscr{U}} &= (C_{\mathscr{U}})^*, \quad C^*|_{\mathscr{Y}} = (C_{\mathscr{Y}})^*, \\ D^*|_{\mathscr{U}} &= (D_{\mathscr{U}})^*, \quad D^*|_{\mathscr{Y}} = (D_{\mathscr{Y}})^*, \end{aligned} \tag{3.60}$$

where recall that $C_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}} C$, $D_{\mathscr{Y}} = \pi_{\mathscr{Y}}^{\mathscr{U}} D$ and so forth.

Proof. We prove (3.59) and (3.60) for C only, as the proof is very similar for D. For $x \in \mathscr{X}$ and $w \in \mathscr{W}$ we see that

$$\begin{aligned} \langle C^\dagger w, x\rangle_{\mathscr{X}} &= [w, Cx]_{\mathscr{W}^*} = -[w, Cx]_{\mathscr{W}} = -\langle Ew, Cx\rangle_{\mathscr{W}} \\ &= -\langle C^* E w, x\rangle_{\mathscr{X}}, \end{aligned}$$

which gives (3.59) by the unicity of the adjoint. To prove (3.60) consider for $u \in \mathscr{U}$ and $x \in \mathscr{X}$

$$\begin{aligned} \langle C_{\mathscr{U}} x, u\rangle_{\mathscr{U}} &= \langle \pi_{\mathscr{U}}^{\mathscr{Y}} Cx, u\rangle_{\mathscr{U}} = \langle Cx, \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{W}} = \langle x, C^* \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{X}} \\ &= \langle x, C^*|_{\mathscr{U}} u\rangle_{\mathscr{X}}, \end{aligned}$$

and again the result follows. The case for $\mathscr{Y}$ is similar. □

Remark 3.4.7. We comment that a signature operator E of an indefinite inner-product space $\mathscr{W}$ is unitary and hence surjective. In particular, the operator E is suitable for considering E-strongly admissible pairs for an output-nulling system with signal space $\mathscr{W}$, see Definition 3.1.11.

Proposition 3.4.8. *Let $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a driving-variable system with indefinite inner-product signal space $(\mathscr{W}, [\cdot,\cdot]_{\mathscr{W}})$ and corresponding signature operator E, and let $\Sigma^* = (\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}}$ denote the dual output-nulling system.*

(i) If the pair $\mathscr{U}, \mathscr{Y}$ is a strongly admissible pair for Σ then $\mathscr{Y}, \mathscr{U}$ is a E-strongly admissible pair for $\Sigma^* = (\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}}$.

Now assume that the pair $\mathscr{U}, \mathscr{Y}$ is strongly admissible for Σ.

(ii) *Let* $(\left[\begin{smallmatrix} \tilde{A}_{\mathrm{D}} & \tilde{C}_{\mathrm{D}} \\ \tilde{B}_{\mathrm{D}} & \tilde{D}_{\mathrm{D}} \end{smallmatrix}\right], \tilde{\mathcal{T}}_{\mathrm{iso}})_{\mathrm{iso}}$ *denote the E-derived $(\mathscr{Y}, \mathscr{U})$ system of Σ^*. Then* $T : \tilde{\mathcal{T}}_{\mathrm{iso}} \to \mathcal{T}^*$ *given by*

$$T \begin{bmatrix} x_* \\ \hline y_* \\ u_* \end{bmatrix} = \begin{bmatrix} x_* \\ E \left[\begin{smallmatrix} -u_* \\ y_* \end{smallmatrix}\right] \end{bmatrix}, \tag{3.61}$$

is an isomorphism.

(iii) *Let* $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ *denote the derived $(\mathscr{U}, \mathscr{Y})$ system of Σ. The following diagram commutes:*

$$\begin{array}{ccc} (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} & \xrightarrow{\text{dual}} & (\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}} \\ {\scriptstyle\text{derived}}\big\downarrow & & \big\downarrow{\scriptstyle\text{derived}} \\ (\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}} & \xrightarrow{\text{dual}} & (\left[\begin{smallmatrix} A_{\mathrm{D}}^* & C_{\mathrm{D}}^* \\ B_{\mathrm{D}}^* & D_{\mathrm{D}}^* \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{iso}}, \end{array}$$

so that $\left[\begin{smallmatrix} A_{\mathrm{D}}^* & C_{\mathrm{D}}^* \\ B_{\mathrm{D}}^* & D_{\mathrm{D}}^* \end{smallmatrix}\right] = \left[\begin{smallmatrix} \tilde{A}_{\mathrm{D}} & \tilde{C}_{\mathrm{D}} \\ \tilde{B}_{\mathrm{D}} & \tilde{D}_{\mathrm{D}} \end{smallmatrix}\right]$.

Proof. (i): From Definition 3.1.11 we are required to prove that $(D^\dagger E)|_{\mathscr{U}} : \mathscr{U} \to \mathscr{V}$ is invertible. By (3.59) and the fact that E is self-adjoint and unitary we have

$$(D^\dagger E)|_{\mathscr{U}} = -(D^* E^2)|_{\mathscr{U}} = -(D^*)|_{\mathscr{U}} = -D^*_{\mathscr{U}}, \tag{3.62}$$

where the last equality is from (3.60). By Definition 3.1.11, $\mathscr{U}, \mathscr{Y}$ a strongly admissible pair for Σ implies that $D_{\mathscr{U}}$ is invertible. Thus from (3.62) we see that that $(D^\dagger E)|_{\mathscr{U}}$ is also invertible.

(ii) : The proof is very similar to that of Theorem 3.1.18. Since we did not give the output-nulling case there we do provide a proof here. The direct sum decomposition $\mathscr{W} = \mathscr{U} \oplus \mathscr{Y}$ and the surjectivity of E implies that any $w_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{W}^*)$ can be written as

$$w_* = -Eu_* + Ey_* = E \begin{bmatrix} -u_* \\ y_* \end{bmatrix}, \tag{3.63}$$

for $u_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{U})$ and $y_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{Y})$. We remark that when deriving input/output pairs u, y from the signal w of an output-nulling trajectory we have chosen to put a minus sign with the component that is the output (which is usually y, compare with (3.27)). In the dual case the input and output spaces interchange and so in (3.61) and (3.63) we have put the minus sign on u_*.

Let x_* and $E\left[\begin{smallmatrix}-u_*\\ y_*\end{smallmatrix}\right]$ denote the components of a trajectory in $\mathcal{T}^*$, so that

$$\begin{aligned}\dot{x}_* &= A^*x - C^\dagger E\begin{bmatrix}-u_*\\ y_*\end{bmatrix} = A^*x_* + (C^\dagger E)|_{\mathscr{U}}u_* - (C^\dagger E)|_{\mathscr{Y}}y_*,\\ 0 &= -B^*x_* + D^\dagger E\begin{bmatrix}-u_*\\ y_*\end{bmatrix} = -B^*x_* - (D^\dagger E)|_{\mathscr{U}}u_* + (D^\dagger E)|_{\mathscr{Y}}y_*\end{aligned} \tag{3.64}$$

We have already seen that $(D^\dagger E)|_{\mathscr{U}}$ is invertible and hence from (3.64) we can obtain an input-state-output relation between y_* and u_*, namely

$$\begin{aligned}\dot{x}_* &= (A^* - (C^\dagger E)|_{\mathscr{U}}(D^\dagger E)|_{\mathscr{U}}^{-1}B^*)x_*\\ &\quad + (-(C^\dagger E)|_{\mathscr{Y}} + (C^\dagger E)|_{\mathscr{U}}(D^\dagger E)|_{\mathscr{U}}^{-1}(D^\dagger E)|_{\mathscr{Y}})y_*,\\ u_* &= -(D^\dagger E)|_{\mathscr{U}}^{-1}B^*x_* + (D^\dagger E)|_{\mathscr{U}}^{-1}(D^\dagger E)|_{\mathscr{Y}}y_*,\end{aligned} \tag{3.65}$$

so that x_* and $\left[\begin{smallmatrix}y_*\\ u_*\end{smallmatrix}\right]$ are the components of a trajectory in $\mathcal{T}^*_{\text{iso}}$. To see the converse we reverse the above steps.

(iii): The bottom route through the diagram gives

$$\begin{aligned}\begin{bmatrix}A & B\\ C & D\end{bmatrix} &\xrightarrow{\text{derived}} \begin{bmatrix}A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1}\\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}\end{bmatrix}\\ &\xrightarrow{\text{dual}} \begin{bmatrix}A^* - (C_{\mathscr{U}})^*D_{\mathscr{U}}^{-*}B^* & (C_{\mathscr{Y}})^* - (C_{\mathscr{U}})^*D_{\mathscr{U}}^{-*}(D_{\mathscr{Y}})^*\\ D_{\mathscr{U}}^{-*}B^* & D_{\mathscr{U}}^{-*}(D_{\mathscr{Y}})^*\end{bmatrix}.\end{aligned} \tag{3.66}$$

The top route through the diagram has effectively already been considered in (3.64) and (3.65). All that remains to verify is that the nodes in (3.65) and (3.66) are the same, but this follows by inspection using (3.59) and (3.60). That the trajectories coincide follows from the definitions of dual trajectories, Theorem 3.1.18 and the relation (3.61). □

3.5 Jointly dissipative systems

Definition 3.5.1. We say that a driving-variable or output-nulling system is jointly signal dissipative if it signal dissipative and its dual is signal dissipative. We say that a driving-variable or output-nulling system is jointly state-signal dissipative if it state-signal dissipative with respect to a positive, self-adjoint operator P and its dual is state-signal dissipative with respect to P^{-1}.

Remark 3.5.2. We remark that we could define a notion of jointly dissipative for an input-state-output system, but do not do so as we shall not require it explicitly. The key point is that the duals of bounded real and positive real systems are again bounded real and positive real respectively.

Lemma 3.5.3. *Let* $\Sigma = (\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *denote a driving-variable system with indefinite inner-product signal space* $\mathscr{W}$ *and let* $\Sigma^* = (\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right], \mathcal{T}^*)_{\mathrm{on}}$ *denote the dual system.*

(i) If Σ *is signal dissipative then* $\operatorname{im} D$ *is non-negative in* $\mathscr{W}$*.*

(ii) If Σ^* *is signal dissipative then* $\ker D^\dagger$ *is non-negative in* $\mathscr{W}^*$*.*

(iii) For $\mathscr{S} \subset \mathscr{W}$*,* $\mathscr{S}$ *is non-positive in* $\mathscr{W}$ *if and only if* $\mathscr{S}$ *is non-negative in* $\mathscr{W}^*$*.*

(iv) $(\operatorname{im} D)^{[\perp]} = \ker D^\dagger$.

Proof. (*i*) Let $v_0 \in \mathscr{V}$ and $x(0) = 0$. Define the constant driving-variable $v(t) = v_0$ for all $t \geq 0$, and let $\left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}_{\mathrm{ext}}$ denote the corresponding trajectory. Note that w is continuous, as v is. The signal dissipativity of Σ implies that

$$0 \leq \int_0^t [w(s), w(s)]_{\mathscr{W}}\, ds = \int_0^t [Cx(s) + Dv(s), Cx(s) + Dv(s)]_{\mathscr{W}}\, ds =: f(t).$$

Clearly $f(0) = 0$ and as f is differentiable, the fundamental theorem of calculus implies that

$$[Cx(0) + Dv(0), Cx(0) + Dv(0)]_{\mathscr{W}} = f'_+(0) := \lim_{h \downarrow 0} \frac{f(h) - f(0)}{h} \geq 0,$$

whence $[Dv_0, Dv_0]_{\mathscr{W}} \geq 0$. As $v_0 \in \mathscr{V}$ was arbitrary we conclude that im D is non-negative in $\mathscr{W}$.

(*ii*) Let $w^0 \in \ker D^\dagger$. We claim that $w^0 = w_*(0)$ for continuous $w_* \in \mathcal{T}^*_{\mathrm{ext}}$. To see this note that as $\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right]$ is an output-nulling node (in particular $D^\dagger$ is surjective) there exists a subspace $\mathscr{Y}$ of $\mathscr{W}$ such that $\ker D^\dagger$ and $\mathscr{Y}$ is a strongly admissible pair (see Lemma 3.1.13). Let $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$ denote the derived $(\ker D^\dagger, \mathscr{Y})$ input-state-output node. From the definition of the derived node (3.25) it follows that $\ker D^\dagger \subseteq \ker D_{\mathrm{D}}$. Now choose a continuous w_1 taking values in $\ker D^\dagger$ with $w_1(0) = w^0$ so that by Theorem 3.1.18, the externally generated input/output trajectory $\left[\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right]$ satisfying

$$\begin{aligned} \dot{x}_* &= A_{\mathrm{D}} x_* + B_D w_1, \\ -w_2 &= C_{\mathrm{D}} x_*, \\ x_*(0) &= 0, \end{aligned} \tag{3.67}$$

is such that $w_* = \left[\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right]$ is an externally generated trajectory of Σ^*. Moreover, since w_1 and x_* are continuous, we see from (3.67) that w_2 is continuous and hence so is w_*. Moreover, $w_2(0) = 0$ and so $w_*(0) = w_1(0) + w_2(0) = w^0$, which proves the above claim.

By signal-dissipativity of Σ^*

$$0 \leq \lim_{t \downarrow 0} \left[\frac{1}{t} \int_0^t [w_*(s), w_*(s)]_{\mathscr{W}^*}\, ds\right] = [w^0, w^0]_{\mathscr{W}^*},$$

as required.

(iii) This is obvious as $v \in \mathscr{S}$ satisfies

$$[v, v]_{\mathscr{W}} \geq 0 \quad \iff \quad [v, v]_{\mathscr{W}^*} = -[v, v]_{\mathscr{W}} \leq 0.$$

(iv) A calculation shows that

$$\begin{aligned} w \in (\operatorname{Im} D)^{[\perp]} &\iff [w, Dv]_{\mathscr{W}} = 0, \quad \forall\, v \in \mathscr{V} \\ &\iff \langle D^{\dagger} w, v\rangle_{\mathscr{V}} = 0, \quad \forall\, v \in \mathscr{V} \\ &\iff D^{\dagger} w = 0, \quad \text{i.e. } w \in \ker D^{\dagger}. \end{aligned}$$

□

Theorem 3.5.4. *Given a jointly signal dissipative driving-variable system, any fundamental decomposition of $\mathscr{W}$ is strongly admissible.*

Proof. Let $\mathscr{W} = \mathscr{W}_+[+]-\mathscr{W}_-$ denote a fundamental decomposition of $\mathscr{W}$. From Lemma 3.5.3 (i), im D is non-negative in $\mathscr{W}$ and combining parts (ii) – (iv) implies that $(\operatorname{im} D)^{[\perp]}$ is non-positive in $\mathscr{W}$. From Lemma 3.2.7 (ii), we see that $\operatorname{im} D$ is maximal non-negative. Thus by Lemma 3.2.7 (i), there is a linear contraction $T : \mathscr{W}_+ \to \mathscr{W}_-$ such that

$$\operatorname{im} D = \mathcal{G}(T),$$

and now it follows from Lemma 3.1.23 (with $T = \tilde{D}$) that the pair $\mathscr{W}_+, \mathscr{W}_-$ is strongly admissible. □

The next result can be seen as a generalisation of the Bounded Real and Positive Real Lemmas to jointly dissipative driving-variable systems.

Theorem 3.5.5. *Given a minimal driving-variable system $\Sigma = \left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T}\right)_{\mathrm{dv}}$, with indefinite inner-product signal space $\mathscr{W}$, the following are equivalent:*

(i) Σ is jointly signal dissipative.

(ii) Σ is jointly state-signal dissipative.

(iii) The signature operator E of $\mathscr{W}$ satisfies $\sigma_+(E) = \dim(\operatorname{im} D)$ and there exists a positive, self-adjoint operator P on $\mathscr{X}$ and operators $M : \mathscr{X} \to \mathscr{U}$, $N : \mathscr{V} \to \mathscr{U}$ satisfying the indefinite KYP Lur'e equations

$$A^*P + PA - C^*EC = -M^*M, \tag{3.68a}$$

$$PB - C^*ED = -M^*N, \tag{3.68b}$$

$$D^*ED = N^*N, \tag{3.68c}$$

where $\mathscr{U}$ is the non-negative part of some fundamental decomposition of $\mathscr{W}$.

In addition, if any of the above hold then there exist positive, self-adjoint solutions P_m, P_M of (3.68) *such that every self-adjoint solution P of* (3.68) *satisfies* $0 < P_m \leq P \leq P_M$. *The extremal operators P_m, P_M are the optimal cost operators of the indefinite optimal control problems, namely:*

$$-\langle P_m x_0, x_0 \rangle_{\mathscr{X}} = \inf_{\substack{\mathcal{T}(x_0) \\ w \in L^2(\mathbb{R}^+; \mathscr{W})}} \int_{\mathbb{R}^+} [w(s), w(s)]_{\mathscr{W}} \, ds, \tag{3.69a}$$

$$-\langle P_M^{-1} x_0, x_0 \rangle_{\mathscr{X}} = \inf_{\substack{\mathcal{T}^*(x_0) \\ w_* \in L^2(\mathbb{R}^+; \mathscr{W}^*)}} \int_{\mathbb{R}^+} [w_*(s), w_*(s)]_{\mathscr{W}^*} \, ds. \tag{3.69b}$$

The minimisation problems (3.69a) *and* (3.69b) *are subject to the driving-variable node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *and dual output-nulling node* $\left[\begin{smallmatrix} A^* & -C^\dagger \\ -B^* & D^\dagger \end{smallmatrix}\right]$ *respectively.*

Proof. *(i)* $\Rightarrow$ *(iii)* : Let $\mathscr{U}, \mathscr{Y}$ denote a fundamental decomposition of $\mathscr{W}$ which by Theorem 3.5.4 is a strongly admissible pair. We recall that with respect to the fundamental decomposition $\mathscr{U}, \mathscr{Y}$ the signature operator E has the block diagonal form

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} : \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix} \to \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix}.$$

Let $\Sigma_{\mathrm{D}} = (\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ denote the derived $(\mathscr{U}, \mathscr{Y})$ system, which by Lemma 3.1.27 is minimal. From Corollary 3.2.8 it follows that

$$\dim(\operatorname{im} D) = \dim \mathscr{U} = \sigma_+(E),$$

and Σ_{D} is bounded real by Theorem 3.3.9. Hence, by the Bounded Real Lemma, there are linear operators $K : \mathscr{X} \to \mathscr{U}$, $W : \mathscr{U} \to \mathscr{U}$ and self-adjoint positive P on $\mathscr{X}$ satisfying the bounded real Lur'e equations (2.10) (with realisation $\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right]$). We prove that the triple (P, K, W) solving (2.10) gives a solution (P, M, N) of (3.68) where

$$M := K + W C_{\mathscr{U}} : \mathscr{X} \to \mathscr{U}, \quad N = W D_{\mathscr{U}} : \mathscr{V} \to \mathscr{U}. \tag{3.70}$$

Expanding first equation (2.10c) we obtain

$$I - D_{\mathrm{D}}^* D_{\mathrm{D}} = I - D_{\mathscr{U}}^{-*} D_{\mathscr{Y}}^* D_{\mathscr{Y}} D_{\mathscr{U}}^{-1} = W^* W \tag{3.71}$$

$$\begin{aligned} \Rightarrow \quad D^* E D = D_{\mathscr{U}}^* D_{\mathscr{U}} - D_{\mathscr{Y}}^* D_{\mathscr{Y}} &= (W D_{\mathscr{U}})^* (W D_{\mathscr{U}}) \\ &= N^* N, \end{aligned} \tag{3.72}$$

where N is as in (3.70). From equations (2.10b) and (3.71) we obtain

$$\begin{aligned} PB - C^*ED &= PB_{\mathrm{D}}D_{\mathscr{U}} - C^*_{\mathscr{U}}D_{\mathscr{U}} + C^*_{\mathscr{Y}}D_{\mathscr{Y}} \\ &= (PB_{\mathrm{D}} + C^*_{\mathrm{D}}D_{\mathrm{D}})D_{\mathscr{U}} - C^*_{\mathscr{U}}W^*WD_{\mathscr{U}} \\ &= -(K + WC_{\mathscr{U}})^*WD_{\mathscr{U}} = -M^*N. \end{aligned} \tag{3.73}$$

Finally, from (2.10a) we infer that

$$\begin{aligned} A^*P + PA - C^*EC =& - K^*K + C^*_{\mathscr{U}}D^{-*}_{\mathscr{U}}B^*P + PBD^{-1}_{\mathscr{U}}C_{\mathscr{U}} \\ &- C^*_{\mathscr{U}}D^{-*}_{\mathscr{U}}D^*_{\mathscr{Y}}D_{\mathscr{Y}}D^{-1}_{\mathscr{U}}C_{\mathscr{U}} - C^*_{\mathscr{U}}C_{\mathscr{U}} \\ =& - (K + WC_{\mathscr{U}})^*(K + WC_{\mathscr{U}}) = -M^*M. \end{aligned} \tag{3.74}$$

Equations (3.72), (3.73) and (3.74) are (3.68a), (3.68b) and (3.68c) respectively.

$(iii) \Rightarrow (ii)$: We can rewrite the equations (3.68) as

$$\begin{aligned} \mathcal{M} := \begin{bmatrix} A^*P + PA - C^*EC & PB - C^*ED \\ B^*P - D^*EC & -D^*ED \end{bmatrix} &= - \begin{bmatrix} M^*M & M^*N \\ N^*M & N^*N \end{bmatrix} \\ &= - \begin{bmatrix} M^* \\ N^* \end{bmatrix} \begin{bmatrix} M & N \end{bmatrix} \\ &\le 0. \end{aligned} \tag{3.75}$$

Given a trajectory $\left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in \mathcal{T}$, let $v \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{V})$ denote a corresponding driving-variable. For $t \ge 0$ from (3.75) we have

$$\int_0^t \left\langle \mathcal{M} \left[\begin{smallmatrix} x(s) \\ v(s) \end{smallmatrix}\right], \left[\begin{smallmatrix} x(s) \\ v(s) \end{smallmatrix}\right] \right\rangle \, ds \le 0,$$

which when unravelled and using (3.12) gives

$$\int_0^t \frac{d}{ds} \langle Px(s), x(s) \rangle_{\mathscr{X}} \le \int_0^t \langle Ew(s), w(s) \rangle_{\mathscr{W}} \, ds. \tag{3.76}$$

Inequality (3.76) is equivalent to

$$\langle Px(t), x(t) \rangle_{\mathscr{X}} - \langle Px_0, x_0 \rangle_{\mathscr{X}} \le \int_0^t [w(s), w(s)]_{\mathscr{W}} \, ds,$$

proving that Σ is state-signal dissipative.

It remains to prove that the dual output-nulling system Σ^* is state-signal dissipative with respect to P^{-1}. By Corollary 3.2.8, $\sigma_+(E) = \dim(\operatorname{im} D)$ implies that any fundamental decomposition $\mathscr{U}$, $\mathscr{Y}$ of $\mathscr{W}$ is strongly admissible. The equations (3.68) collapse to the bounded real Lur'e equations for the derived $(\mathscr{U}, \mathscr{Y})$ system, which thus have a positive, self-adjoint solution P. We infer that Σ_{D} is bounded real, and hence so is the dual input-state-output system Σ_{D}^*. As stated in the discussion in Section 2.2, it follows that the dual bounded real Lur'e equations (2.15) have positive, self-adjoint solution P^{-1}. Hence by part (iii) of the Bounded Real Lemma applied to Σ_{D}^* we see that for trajectories of Σ_{D}^* with state $x_* \in C(\mathbb{R}^+;\mathscr{X})$, input $y_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{Y})$ and output $u_* \in L^2_{\mathrm{loc}}(\mathbb{R}^+;\mathscr{U})$ and for $t \geq 0$

$$\langle P^{-1}x_*(t), x_*(t)\rangle_{\mathscr{X}} - \langle P^{-1}x_*(0), x_*(0)\rangle_{\mathscr{X}} \leq \int_0^t \|y_*(s)\|^2_{\mathscr{Y}} - \|u_*(s)\|^2_{\mathscr{U}}\, ds. \tag{3.77}$$

However, from Proposition 3.4.8 the trajectories of Σ^* and Σ_D^* are related by (3.61). In particular, every signal w_* of Σ^* (with state x_*) can be decomposed as

$$E\begin{bmatrix} -u_* \\ y_* \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} -u_* \\ y_* \end{bmatrix} = \begin{bmatrix} -u_* \\ -y_* \end{bmatrix},$$

where y_*, u_* is an input/output pair for Σ_D^* also with state x_*. Therefore,

$$\begin{aligned}
\int_0^t [w_*(s), w_*(s)]_{\mathscr{W}^*}\, ds &= \int_0^t -\langle w_*(s), Ew_*(s)\rangle_{\mathscr{W}}\, ds \\
&= \int_0^t -\left\langle \begin{bmatrix} -u_*(s) \\ -y_*(s) \end{bmatrix}, \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} -u_*(s) \\ -y_*(s) \end{bmatrix} \right\rangle_{\mathscr{U}\times\mathscr{Y}} ds \\
&= \int_0^t \|y_*(s)\|^2_{\mathscr{Y}} - \|u_*(s)\|^2_{\mathscr{Y}}\, ds.
\end{aligned} \tag{3.78}$$

Combining (3.77) and (3.78) we conclude that Σ^* is state-signal dissipative with respect to P^{-1}, as required.

$(ii) \Rightarrow (i)$: This implication is trivial.

Finally, suppose any (hence all) of (i)-(iii) above hold. Choose a fundamental decomposition of the signal space. This decomposition is strongly admissible by Theorem 3.5.4 and provides a bounded real derived input-state-output system by Theorem 3.3.9. From the proof of $(i) \Rightarrow (ii)$ above it follows that there is a one to one correspondence between triples (P, K, W) solving (2.10) and triples (P, M, N) solving (3.68). By the Bounded Real Lemma there exist extremal positive self-adjoint operators P_m, P_M to (2.10), which are therefore extremal positive self-adjoint solutions of (3.68). Furthermore from the Bounded Real Lemma we see that for any $x_0 \in \mathscr{X}$

$$-\langle P_m x_0, x_0\rangle_{\mathscr{X}} = \inf_{u\in L^2(\mathbb{R}^+;\mathscr{U})} \int_{\mathbb{R}^+} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}}\, ds,$$

subject to the input-state-output node $\begin{bmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{bmatrix}$. We can rewrite this as

$$\begin{aligned} -\langle P_m x_0, x_0 \rangle_{\mathscr{X}} &= \inf_{\substack{\mathcal{T}(x_0) \\ u \in L^2(\mathbb{R}^+;\mathscr{U})}} \int_{\mathbb{R}^+} \|u(s)\|^2_{\mathscr{U}} - \|y(s)\|^2_{\mathscr{Y}} \, ds, \\ &= \inf_{\substack{\mathcal{T}(x_0) \\ w \in L^2(\mathbb{R}^+;\mathscr{W})}} \int_{\mathbb{R}^+} [w(s), w(s)]_{\mathscr{W}} \, ds, \end{aligned}$$

since the trajectories from x_0 of Σ and Σ_{D} are precisely the same by Theorem 3.1.18. The dual case is similar, only now starting from the fact that

$$-\langle P_M^{-1} x_0, x_0 \rangle_{\mathscr{X}} = \inf_{y_* \in L^2(\mathbb{R}^+;\mathscr{Y})} \int_{\mathbb{R}^+} \|y_*(s)\|^2_{\mathscr{Y}} - \|u_*(s)\|^2_{\mathscr{U}} \, ds,$$

subject to the dual input-state-output node $\begin{bmatrix} A^*_{\mathrm{D}} & C^*_{\mathrm{D}} \\ B^*_{\mathrm{D}} & D^*_{\mathrm{D}} \end{bmatrix}$. These observations establish (3.69a) and (3.69b). □

Remark 3.5.6.

1. By Theorem 3.5.5 above the two notions of dissipativity are equivalent for minimal jointly dissipative driving-variable systems. There is therefore no ambiguity in calling such systems simply jointly dissipative.

2. The assumption that $\sigma_+(E) = \dim(\operatorname{im} D)$ is essential to prove the implication $(iii) \Rightarrow (ii)$. Although existence of a positive, self-adjoint solution P to the indefinite KYP Lur'e equations (3.68) does imply state-signal dissipativity of the system, it does not necessarily imply state-signal dissipativity of the dual system. The assumption that $\sigma_+(E) = \dim(\operatorname{im} D)$ also features in behavorial theory, and is a property referred to as *liveness* (see [102, p.56]). As we have already seen, for our purposes it ensures that the dual of a system is dissipative when the system itself is dissipative.

3.6 Dissipative balanced approximations

In this section we define the dissipative balanced truncation of a minimal jointly dissipative driving-variable system. The aim is to derive a gap metric error bound for dissipative balanced truncation, as well as to generalise bounded real and positive real balanced truncation (so that we see them as special cases). To do these tasks we make use of the Indefinite KYP Lemma, Theorem 3.5.5.

3.6.1 Dissipative balanced truncation

Definition 3.6.1. A minimal, jointly dissipative driving-variable system is called dissipative balanced or in dissipative balanced co-ordinates if

$$P_m = P_M^{-1} =: \Pi, \tag{3.79}$$

where P_m and P_M are the extremal solutions of the KYP Lur'e equations (3.68). We denote by $(\sigma_i^2)_{i=1}^m$ the eigenvalues of $P_m P_M^{-1}$, ordered in decreasing magnitude, each with multiplicity r_i. We call $(\sigma_i)_{i=1}^m$ the dissipative characteristic values.

Proposition 3.6.2. *Let* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *denote a minimal jointly dissipative driving-variable system. Then there exists an invertible operator* $T : \mathscr{X} \to \mathscr{X}$ *such that the similarity transformed system* $(\left[\begin{smallmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *is dissipative balanced. We call such a* T *a dissipative balancing transformation.*

Proof. This is an application of [3, Lemma 7.3]. □

Definition 3.6.3. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ denote a minimal jointly dissipative driving-variable system in dissipative balanced co-ordinates. Let $(\sigma_i)_{i=1}^m$ denote the dissipative characteristic values, with multiplicities r_i. For $r < m$ let $\mathscr{X}_r$ denote the sum of the first r eigenspaces of Π, with corresponding orthogonal projection $P_{\mathscr{X}_r}$. Define the operators

$$\Pi_1 = P_{\mathscr{X}_r}\Pi|_{\mathscr{X}_r}, \qquad A_{11} = P_{\mathscr{X}_r}A|_{\mathscr{X}_r}, \qquad B_1 = P_{\mathscr{X}_r}B, \qquad C_1 = C|_{\mathscr{X}_r}. \tag{3.80}$$

Let $\mathcal{T}_r$ denote the trajectories corresponding to the driving-variable node $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$. We call the driving-variable system $(\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right], \mathcal{T}_r)_{\mathrm{dv}}$ the dissipative balanced truncation (of order $\sum_{i=1}^r r_i$) of $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$.

Definition 3.6.4. Let $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ denote a minimal input-state-output system that is signal dissipative with respect to an indefinite inner-product with signature operator E that satisfies $\sigma_+(E) = \dim \mathscr{U}$. We say that the input-state-output system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ is dissipative balanced if the jointly dissipative driving-variable system

$$\left(\left[\begin{array}{c|c} A & B \\ \hline 0 & I \\ C & D \end{array}\right], \mathcal{T} \right)_{\mathrm{dv}},$$

is dissipative balanced in the sense of Definition 3.6.1. We call the input-state-output system corresponding to the node $\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right]$, where A_{11}, B_1 and C_1 are as in (3.80) the dissipative balanced truncation (of order $\sum_{i=1}^r r_i$) of $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$.

Lemma 3.6.5. *Let* $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *denote a driving-variable system with strongly admissible pair* $\mathscr{U}, \mathscr{Y}$ *and let* T *denote an invertible operator on the state space. Letting* $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ *denote the derived* $(\mathscr{U}, \mathscr{Y})$ *system the following diagram commutes*

$$\begin{tikzcd}
(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} \arrow[r, "\text{transform}"] \arrow[d, "\text{derived}"'] & (\left[\begin{smallmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} \arrow[d, "\text{derived}"] \\
(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}} \arrow[r, "\text{transform}"] & (\left[\begin{smallmatrix} T^{-1}A_{\mathrm{D}}T & T^{-1}B_{\mathrm{D}} \\ C_{\mathrm{D}}T & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}.
\end{tikzcd}$$

If additionally $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}}$ *is minimal, jointly dissipative and dissipative balanced then the following diagram commutes*

$$\begin{tikzcd}
(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} \arrow[r, "\text{truncate}"] \arrow[d, "\text{derived}"'] & (\left[\begin{smallmatrix} A_{11} & B_1 \\ C_1 & D \end{smallmatrix}\right], \mathcal{T}_r)_{\mathrm{dv}} \arrow[d, "\text{derived}"] \\
(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}} \arrow[r, "\text{truncate}"] & (\left[\begin{smallmatrix} (A_{\mathrm{D}})_{11} & (B_{\mathrm{D}})_1 \\ (C_{\mathrm{D}})_1 & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T}_r)_{\mathrm{iso}}.
\end{tikzcd}$$

As such the derived $(\mathscr{U}, \mathscr{Y})$ *system of a dissipative balanced truncation driving-variable system is the same as taking the derived* $(\mathscr{U}, \mathscr{Y})$ *system of the original driving-variable system and then dissipative balancing and truncating the resulting input-state-output system.*

Proof. Consider the first diagram. The bottom route through the diagram states that

$$\begin{aligned}
\begin{bmatrix} A & B \\ C & D \end{bmatrix} &\xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} \\
&\xrightarrow{\text{transform}} \begin{bmatrix} T^{-1}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})T & T^{-1}BD_{\mathscr{U}}^{-1} \\ (C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}})T & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}.
\end{aligned} \tag{3.81}$$

The top route through the diagram states that

$$\begin{aligned}
\begin{bmatrix} A & B \\ C & D \end{bmatrix} &\xrightarrow{\text{transform}} \begin{bmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{bmatrix} \\
&\xrightarrow{\text{derived}} \begin{bmatrix} T^{-1}AT - T^{-1}BD_{\mathscr{U}}^{-1}(CT)_{\mathscr{U}} & T^{-1}BD_{\mathscr{U}}^{-1} \\ (CT)_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}(C_{\mathscr{U}})T & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}.
\end{aligned} \tag{3.82}$$

That (3.81) and (3.82) are the same follows by inspection and the fact that $(CT)_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}}CT = C_{\mathscr{U}}T$ (and similarly for $\mathscr{Y}$ instead of $\mathscr{U}$).

We now prove that the second diagram commutes. The bottom route gives

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}$$
$$\xrightarrow{\text{truncate}} \begin{bmatrix} P_{\mathscr{X}_r}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})|_{\mathscr{X}_r} & P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1} \\ (C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}})|_{\mathscr{X}_r} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{3.83}$$

The top route through the diagram states that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{truncate}} \begin{bmatrix} P_{\mathscr{X}_r}A|_{\mathscr{X}_r} & P_{\mathscr{X}_r}B \\ C|_{\mathscr{X}_r} & D \end{bmatrix}$$
$$\xrightarrow{\text{derived}} \begin{bmatrix} P_{\mathscr{X}_r}A|_{\mathscr{X}_r} - P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1}(C|_{\mathscr{X}_r})_{\mathscr{U}} & P_{\mathscr{X}_r}BD_{\mathscr{U}}^{-1} \\ (C|_{\mathscr{X}_r})_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}(C|_{\mathscr{X}_r})_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix}. \tag{3.84}$$

That (3.83) and (3.84) are the same again follows by inspection, the facts that restriction and projection are linear and that $(C|_{\mathscr{X}_r})_{\mathscr{U}} = \pi_{\mathscr{U}}^{\mathscr{Y}} C|_{\mathscr{X}_r} = (C_{\mathscr{U}})|_{\mathscr{X}_r}$ (and similarly for $\mathscr{Y}$ instead of $\mathscr{U}$). ☐

Corollary 3.6.6. *Let Σ denote a minimal jointly dissipative driving-variable system, let $\mathscr{U}$, $\mathscr{Y}$ denote a strongly admissible pair and let Σ_{D} denote the derived $(\mathscr{U}, \mathscr{Y})$ system.*

(i) If Σ_{D} is bounded real (positive real) then the dissipative characteristic values of Σ are precisely the bounded real singular values (positive real singular values) of Σ_{D}.

(ii) If Σ is dissipative balanced and Σ_{D} is bounded real (positive real) then Σ_{D} is bounded real balanced (positive real balanced).

(iii) Let Σ_r denote the dissipative balanced truncation of order $\sum_{i=1}^{r} r_i$ (using the notation of Definition 3.6.3). If Σ_{D} is bounded real (positive real) then the $(\mathscr{U}, \mathscr{Y})$ derived system of Σ_r is the bounded real balanced truncation (positive real balanced truncation) of order $\sum_{i=1}^{r} r_i$ of Σ_{D}.

Proof. Claims (*i*) and (*ii*) follow from the indefinite KYP Lemma and the Bounded Real Lemma (Positive Real Lemma). Specifically the extremal solutions P_m and P_M of the KYP Lur'e equations (3.68) of Σ *are* the extremal solutions of the bounded real Lur'e equations (2.10) (positive real Lur'e equations (2.24)) of the $(\mathscr{U}, \mathscr{Y})$ derived system. Claim (*iii*) then follows from claims (*i*) and (*ii*) and the commuting diagrams in Lemma 3.6.5. ☐

Using the above connections to bounded real and positive real balanced truncation, it follows that dissipative balanced truncation preserves joint dissipativity.

Corollary 3.6.7. *Let Σ denote a minimal jointly dissipative driving-variable system. Then for every r the dissipative balanced truncation Σ_r is jointly dissipative.*

Proof. This follows from Theorem 3.5.5 and Corollary 3.6.6. □

We are now in position to state and prove our main result, which is an error bound in the gap metric for dissipative balanced truncation.

Theorem 3.6.8. *Given a minimal jointly dissipative driving-variable system Σ let $(\sigma_i)_{i=1}^m$ denote the dissipative characteristic values and for $r < m$ let Σ_r denote the dissipative balanced truncation of Σ. Then*

$$\hat{\delta}(\Sigma, \Sigma_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i. \tag{3.85}$$

Proof. We make use of Theorem 2.2.7, an error bound for bounded real input-state-output systems. Choose $r < m$ and a fundamental decomposition $\mathscr{U}, \mathscr{Y}$ of the signal space, which is strongly admissible by Theorem 3.5.4. The derived $(\mathscr{U}, \mathscr{Y})$ system is bounded real by Theorem 3.3.9. Denote the transfer function and input-output map of this system by G and $\mathfrak{D}_G$ respectively. By Corollary 3.6.6 (*i*) the bounded real singular values of G are precisely the dissipative characteristic values of Σ. By Corollary 3.6.6 (*iii*) the bounded real balanced truncation, with transfer function G_r and input-output map $\mathfrak{D}_{G_r}$, is the derived $(\mathscr{U}, \mathscr{Y})$ system of Σ_r. The error bound (2.18)

$$\|G - G_r\|_{H^\infty} \leq 2 \sum_{i=r+1}^{m} \sigma_i,$$

from Theorem 2.2.7 then applies. Let $\mathcal{S}$ and $\mathcal{S}_r$ denote the sets of stable externally generated trajectories of Σ and Σ_r respectively. Since G is bounded real, it belongs to H^∞ and hence the input-output map $\mathfrak{D}_G$ is bounded $L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$. In this instance the conclusion of Corollary 3.1.20 can be strengthened to

$$\mathcal{S} = \mathcal{G}(\mathfrak{D}_G).$$

The same relation holds for the dissipative balanced truncation and the bounded real balanced truncation of the derived system (for the same reasons), namely

$$\mathcal{S}_r = \mathcal{G}(\mathfrak{D}_{G_r}).$$

Therefore

$$\hat{\delta}(\Sigma, \Sigma_r) := \hat{\delta}(\mathcal{S}, \mathcal{S}_r) = \hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_r})) = \hat{\delta}(\mathfrak{D}_G, \mathfrak{D}_{G_r}). \tag{3.86}$$

We also require the well-known equality (see for example [94])

$$\|\mathfrak{D}_G - \mathfrak{D}_{G_r}\| = \|G - G_r\|_{H^\infty} . \tag{3.87}$$

Combining the bounds (2.18) and (3.39) with the equalities (3.86) and (3.87) yields the desired bound (3.85). $\square$

The above bound holds for any strongly admissible decomposition into inputs and outputs (since the gap metric is independent of such a splitting), and in particular for the positive real case. Therefore we obtain a gap metric error bound for stable positive real balanced truncation and, by using the equivalence between the H^∞ norm and the gap metric for stable systems, we get an H^∞ error bound for positive real systems. The gap metric error bound (3.88) has been independently established by Timo Reis [69].

Corollary 3.6.9. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real rational transfer function with positive real singular values $(\sigma_i)_{i=1}^m$ and for $r < m$ let J_r denote the positive real balanced truncation. Then the following bounds hold,*

$$\hat{\delta}(J, J_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i, \tag{3.88}$$

and

$$\begin{aligned}\|J - J_r\|_{H^\infty} \leq 2 \min \Big\{ &(1 + \|J\|_{H^\infty}^2)(1 + \|J_r\|_{H^\infty}), \\ &(1 + \|J\|_{H^\infty})(1 + \|J_r\|_{H^\infty}^2) \Big\} \sum_{i=r+1}^{m} \sigma_i. \end{aligned} \tag{3.89}$$

In inequality (3.88) *we are abusing notation by writing $\hat{\delta}(J, J_r) = \hat{\delta}(\mathfrak{D}_J, \mathfrak{D}_{J_r})$, where $\mathfrak{D}_J$ and $\mathfrak{D}_{J_r}$ are the input-output maps corresponding to J and J_r respectively.*

Remark 3.6.10. The H^∞ error bound (3.89) is not an *a priori* error bound, as it requires $\|J_r\|_{H^\infty}$, thus limiting its usefulness in practise.

Proof of Corollary 3.6.9: Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal realisation of J, from which we build a driving-variable system Σ as in Proposition 3.1.21, which is minimal and dissipative with respect to the indefinite inner-product induced by $E := \left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]$ by Proposition 3.3.7. Since the dual input-state-output system with node $\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$ is positive real, by Proposition 3.4.8 we see that Σ is jointly dissipative. By Corollary 3.6.6 (i) the dissipative characteristic values of Σ are precisely the positive real singular values of J and by part (iii) of that result the $(\mathscr{U}, \mathscr{U})$ derived dissipative balanced truncation of Σ is the positive real balanced truncation of J. The gap metric error bound for $\hat{\delta}(J, J_r)$ now follows from Theorem 3.6.8 and Theorem 3.1.18.

To prove the H^∞ bound we use the equivalence of the gap metric restricted to bounded, linear operators and the operator norm. Let $\mathfrak{D}_J$ and $\mathfrak{D}_{J_r}$ denote the input-output maps for J and its positive real balanced truncation J_r respectively. The input-output maps are bounded as $J, J_r \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$. Let $\mathfrak{A} = I + \mathfrak{D}_J^* \mathfrak{D}_J$, and define $\mathfrak{A}_r$ similarly so that by Partington [64, p. 32] the orthogonal projection $P_{\mathcal{G}(\mathfrak{D}_J)}$ onto $\mathcal{G}(\mathfrak{D}_J)$ is given by

$$P_{\mathcal{G}(\mathfrak{D}_J)} = \begin{bmatrix} I \\ \mathfrak{D}_J \end{bmatrix} \begin{bmatrix} I + \mathfrak{D}_J^* \mathfrak{D}_J \end{bmatrix}^{-1} \begin{bmatrix} I & \mathfrak{D}_J^* \end{bmatrix} = \begin{bmatrix} \mathfrak{A}^{-1} & \mathfrak{A}^{-1}\mathfrak{D}_J^* \\ \mathfrak{D}_J \mathfrak{A}^{-1} & \mathfrak{D}_J \mathfrak{A}^{-1} \mathfrak{D}_J^* \end{bmatrix}.$$

Thus

$$\begin{aligned} \hat{\delta}(\mathcal{G}(\mathfrak{D}_J), \mathcal{G}(\mathfrak{D}_{J_r})) &= \left\| P_{\mathcal{G}(\mathfrak{D}_J)} - P_{\mathcal{G}(\mathfrak{D}_{J_r})} \right\| \\ &= \left\| \begin{bmatrix} \mathfrak{A}^{-1} - \mathfrak{A}_r^{-1} & \mathfrak{A}^{-1}\mathfrak{D}_J^* - \mathfrak{A}_r^{-1}\mathfrak{D}_{J_r}^* \\ \mathfrak{D}_J\mathfrak{A}^{-1} - \mathfrak{D}_{J_r}\mathfrak{A}_r^{-1} & \mathfrak{D}\mathfrak{A}^{-1}\mathfrak{D}_J^* - \mathfrak{D}_{J_r}\mathfrak{A}_r^{-1}\mathfrak{D}_{J_r}^* \end{bmatrix} \right\| \\ &\geq \begin{cases} \left\| \mathfrak{A}^{-1} - \mathfrak{A}_r^{-1} \right\|. \\ \left\| \mathfrak{D}_J \mathfrak{A}^{-1} - \mathfrak{D}_{J_r} \mathfrak{A}_r^{-1} \right\|. \end{cases} \end{aligned}$$

We calculate

$$\begin{aligned} \|\mathfrak{D}_J - \mathfrak{D}_{J_r}\| &= \left\| \mathfrak{D}_J \mathfrak{A}^{-1}\mathfrak{A} - \mathfrak{D}_{J_r}\mathfrak{A}^{-1}\mathfrak{A} \right\| \leq \|\mathfrak{A}\| . \left\| \mathfrak{D}_J\mathfrak{A}^{-1} - \mathfrak{D}_{J_r}\mathfrak{A}^{-1} \right\| \qquad (3.90) \\ &\leq \|\mathfrak{A}\| . \left(\left\| \mathfrak{D}_J\mathfrak{A}^{-1} - \mathfrak{D}_{J_r}\mathfrak{A}_r^{-1} \right\| + \left\| \mathfrak{D}_{J_r}\mathfrak{A}_r^{-1} - \mathfrak{D}_{J_r}\mathfrak{A}^{-1} \right\| \right) \\ &\leq \|\mathfrak{A}\|.(1 + \|\mathfrak{D}_{J_r}\|) \left\| P_{\mathcal{G}(\mathfrak{D}_J)} - P_{\mathcal{G}(\mathfrak{D}_{J_r})} \right\| \\ &\leq (1 + \|\mathfrak{D}_J\|^2)(1 + \|\mathfrak{D}_{J_r}\|)\hat{\delta}(\mathcal{G}(\mathfrak{D}_J), \mathcal{G}(\mathfrak{D}_{J_r})). \qquad (3.91) \end{aligned}$$

Using the gap metric error bound (3.85) we see that from inequality (3.91)

$$\|J - J_r\|_{H^\infty} \leq 2(1 + \|J\|_{H^\infty}^2)(1 + \|J_r\|_{H^\infty}) \sum_{i=k+1}^{m} \sigma_i. \qquad (3.92)$$

Finally we note that in the above proof we could have started with $I = \mathfrak{A}_r^{-1}\mathfrak{A}_r$ instead of $I = \mathfrak{A}^{-1}\mathfrak{A}$ in equation (3.90). In this case we get the alternative error bound in (3.89), namely

$$\|J - J_r\|_{H^\infty} \leq 2(1 + \|J\|_{H^\infty})(1 + \|J_r\|_{H^\infty}^2) \sum_{i=k+1}^{m} \sigma_i.$$

□

Remark 3.6.11. In some circumstances the error bound (3.89) can be made more conservative, with the advantage of not including $\|J_r\|_{H^\infty}$ in the right hand side. Specifically, using the notation of Corollary 3.6.9, if r and J are such that

$$2(1+\|J\|_{H^\infty}^2)\sum_{j=r+1}^{m}\sigma_j<1,$$

then

$$\|J-J_r\|_{H^\infty}\leq\frac{(1+\|J\|_{H^\infty})(1+\|J\|_{H^\infty}^2)\sum_{j=r+1}^{m}\sigma_j}{1-2(1+\|J\|_{H^\infty}^2)\sum_{j=r+1}^{m}\sigma_j}. \tag{3.93}$$

The bound (3.93) follows from (3.92) by setting $\alpha=\|J-J_r\|_{H^\infty}$, using

$$\|J_r\|_{H^\infty}\leq\alpha+\|J\|_{H^\infty},$$

and turning the resulting expression into a bound for α.

3.6.2 Singular perturbation approximation

So far we have considered model reduction by direct truncation. In this section we demonstrate that singular perturbation approximation is often another suitable method for model reduction of a driving-variable system.

Definition 3.6.12. Let $(\left[\begin{smallmatrix}A&B\\C&D\end{smallmatrix}\right],\mathcal{T})_{\mathrm{dv}}$ denote a minimal jointly dissipative driving-variable system in dissipative balanced co-ordinates. Let $(\sigma_i)_{i=1}^m$ denote the dissipative characteristic values, each with multiplicity r_i. For $s<m$ let $\mathscr{X}_s$ and $\mathscr{Z}_s$ denote the sum of the first s and last $m-s$ eigenspaces of Π respectively, with respective orthogonal projections $P_{\mathscr{X}_s}$ and $P_{\mathscr{Z}_s}$. Then with respect to the orthogonal decomposition $\mathscr{X}=\mathscr{X}_s\oplus\mathscr{Z}_s$, the operators A,B,C and Π split as

$$\Pi=\begin{bmatrix}P_{\mathscr{X}_s}\Pi|_{\mathscr{X}_s}&0\\0&P_{\mathscr{Z}_s}\Pi|_{\mathscr{Z}_s}\end{bmatrix}=\begin{bmatrix}\Pi_1&0\\0&\Pi_2\end{bmatrix},\qquad B=\begin{bmatrix}B_1\\B_2\end{bmatrix},$$
$$A=\begin{bmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{bmatrix},\qquad C=\begin{bmatrix}C_1&C_2\end{bmatrix}.$$

Assuming that A_{22}^{-1} exists define

$$\begin{aligned}A_{\mathrm{s}}&=A_{11}-A_{12}A_{22}^{-1}A_{21}, & B_{\mathrm{s}}&=B_1-A_{12}A_{22}^{-1}B_2,\\ C_{\mathrm{s}}&=C_1-C_2A_{22}^{-1}A_{21}, & D_{\mathrm{s}}&=D-C_2A_{22}^{-1}B_2.\end{aligned} \tag{3.94}$$

Let $\mathcal{T}_s$ denote the trajectories corresponding to the driving-variable node $\left[\begin{smallmatrix}A_{\mathrm{s}}&B_{\mathrm{s}}\\C_{\mathrm{s}}&D_{\mathrm{s}}\end{smallmatrix}\right]$. We call the driving-variable system $(\left[\begin{smallmatrix}A_{\mathrm{s}}&B_{\mathrm{s}}\\C_{\mathrm{s}}&D_{\mathrm{s}}\end{smallmatrix}\right],\mathcal{T}_{\mathrm{s}})_{\mathrm{dv}}$ the (dissipative) singular perturbation approximation (of order $\sum_{i=1}^s r_i$) of $(\left[\begin{smallmatrix}A&B\\C&D\end{smallmatrix}\right],\mathcal{T})_{\mathrm{dv}}$.

Lemma 3.6.13. *Let* $(\left[\begin{smallmatrix}A&B\\C&D\end{smallmatrix}\right],\mathcal{T})_{\mathrm{dv}}$ *denote a minimal jointly dissipative driving-variable system in dissipative balanced co-ordinates with strongly admissible pair* $\mathscr{U},\mathscr{Y}$ *and* *let* $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$ *denote the derived* $(\mathscr{U}, \mathscr{Y})$ *system. If the singular perturbation approximation* $(\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s)_{\mathrm{dv}}$ *exists then the following diagram commutes*

$$\begin{tikzcd}
(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{dv}} \arrow[r, "\mathrm{spa}"] \arrow[d, "\mathrm{derived}"'] & (\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s)_{\mathrm{dv}} \arrow[d, "\mathrm{derived}"] \\
(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}} \arrow[r, "\mathrm{spa}"] & (\left[\begin{smallmatrix} (A_{\mathrm{D}})_{\mathrm{s}} & (B_{\mathrm{D}})_{\mathrm{s}} \\ (C_{\mathrm{D}})_{\mathrm{s}} & (D_{\mathrm{D}})_{\mathrm{s}} \end{smallmatrix}\right])_{\mathrm{iso}}
\end{tikzcd}$$

(here spa *denotes taking the singular perturbation approximation) provided that the inverses*

$$(A_{22} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2})^{-1}, \tag{3.95}$$

$$\text{and} \quad [\pi_{\mathscr{U}}^{\mathscr{Y}}(D - C_2 A_{22}^{-1} B_2)]^{-1}, \tag{3.96}$$

exist, where $C_{\mathscr{U},2} = \pi_{\mathscr{U}}^{\mathscr{Y}} C_2$.

Proof. The proof is a rather long, but elementary, series of calculations which appear in Appendix A. The requirement that the inverse in (3.95) exists ensures that we can take the singular perturbation approximation of $(\left[\begin{smallmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{smallmatrix}\right], \mathcal{T})_{\mathrm{iso}}$. The existence of the inverse in (3.96) implies that $\mathscr{U}, \mathscr{Y}$ is a strongly admissible pair for $(\left[\begin{smallmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{smallmatrix}\right], \mathcal{T}_s)_{\mathrm{dv}}$. □

Theorem 3.6.14. *Given a minimal jointly dissipative driving-variable system* Σ *let* $(\sigma_i)_{i=1}^m$ *denote the dissipative characteristic values and for* $s < m$ *assume that the singular perturbation approximation* Σ_s *exists. If the inverses in* (3.95) *and* (3.96) *exist for* $\mathscr{U}, \mathscr{Y}$ *a fundamental decomposition of* $\mathscr{W}$ *then*

$$\hat{\delta}(\Sigma, \Sigma_s) \leq 2 \sum_{i=s+1}^{m} \sigma_i. \tag{3.97}$$

Proof. The proof is identical to that of Theorem 3.6.8, only instead appealing to [53, Theorem 3] for the error bound

$$\|G - G_s\|_{H^\infty} \leq 2 \sum_{i=s+1}^{m} \sigma_i,$$

where G and G_s are the transfer function of the derived $(\mathscr{U}, \mathscr{Y})$ system and its singular perturbation approximation respectively. Lemma 3.6.13 ensures that $\hat{\delta}(\Sigma, \Sigma_s) = \hat{\delta}(\mathfrak{D}_G, \mathfrak{D}_{G_s})$, where $\mathfrak{D}_G$ and $\mathfrak{D}_{G_s}$ are the input-output maps of G and G_s respectively. □

Similar adaptations can be made to Corollary 3.6.9, to give the following error bounds of the difference of the transfer function of a minimal positive real input-state-output system and its singular perturbation approximation.

Corollary 3.6.15. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a positive real rational transfer function with positive real singular values $(\sigma_i)_{i=1}^m$ and assume for that for $s < m$ the singular perturbation approximation J_s of J exists. Then the following bound holds,*

$$\hat{\delta}(J, J_s) \leq 2 \sum_{i=s+1}^{m} \sigma_i, \tag{3.98}$$

and

$$\begin{aligned} \|J - J_s\|_{H^\infty} \leq 2 \min \Big\{ &(1 + \|J\|_{H^\infty}^2)(1 + \|J_s\|_{H^\infty}), \\ &(1 + \|J\|_{H^\infty})(1 + \|J_s\|_{H^\infty}^2) \Big\} \sum_{i=s+1}^{m} \sigma_i. \end{aligned} \tag{3.99}$$

Proof. The proof is identical to that of Corollary 3.6.9, only with changing $\mathfrak{D}_r$ to $\mathfrak{D}_s$, the input-output map of J_s. We also use the gap metric error bound (3.97) instead of (3.85). □

3.6.3 Some remarks on improper rational functions

So far in this chapter we have considered dissipative state-signal systems with a view to deriving classical input-state-output systems, that is input-output relations described by proper rational transfer functions. We have used the terms strongly admissible pairs and strongly derived systems to describe such situations. It is possible, however, to relate the signals of a state-signal system to a wider class of input-output systems by considering a weaker notion of admissibility. In this section we present a few results in this direction.

Definition 3.6.16. Let $\mathscr{U}$ and $\mathscr{Y}$ denote finite-dimensional Hilbert spaces. Given a rational $B(\mathscr{U}, \mathscr{Y})$-valued function G, we define the set of stable trajectories corresponding to G, $\mathcal{S}_G$, by

$$\mathcal{S}_G = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) \; : \; \exists\, \alpha \in \mathbb{R} \text{ such that } \hat{y}(s) = G(s)\hat{u}(s), \, \forall\, s \in \mathbb{C}_\alpha^+ \right\}. \tag{3.100}$$

We call G the transfer function between u and y and the transfer function G together with trajectories $\mathcal{S}_G$ an input-output system.

Example 3.6.17. The scalar rational function $G(s) = s$ has corresponding set of stable trajectories

$$\mathcal{S}_G = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \begin{bmatrix} W^{1,2}(\mathbb{R}^+; \mathscr{U}) \\ L^2(\mathbb{R}^+; \mathscr{Y}) \end{bmatrix} \; : \; u(0) = 0, \quad y = \dot{u} \right\},$$

where $W^{1,2}$ is the usual Sobolev space.

Remark 3.6.18. An input-state-output system of Definition 3.1.2 is an input-output system, with the same transfer function. If $\mathcal{S}$ and G denote the set of externally generated trajectories and transfer function respectively, then $\mathcal{S} = \mathcal{S}_G$.

Remark 3.6.19. In this section, for a driving-variable system, we let $\mathcal{S}$ denote the set of stable signals (that is $w \in L^2(\mathbb{R}^+; \mathscr{W})$) such that the corresponding driving-variable v has a Laplace transform (on some right-half plane). In Definition 3.1.6 we only required that $v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V})$. Certainly if

$$v \in \bigcup_{\alpha \geq 0} L^2_\alpha(\mathbb{R}^+; \mathscr{V}),$$

then $v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V})$ and v has a Laplace transform. Here $x \in L^2_\alpha(\mathbb{R}^+, \mathscr{V})$ if and only if $t \mapsto \mathrm{e}^{-\alpha t} x(t) \in L^2(\mathbb{R}^+; \mathscr{V})$.

Proposition 3.6.20. *If G is a rational $B(\mathscr{U}, \mathscr{Y})$ valued function then there exists a driving-variable system with signal space $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ and set of stable trajectories $\mathcal{S}$ such that*

$$\mathcal{S}_G = \mathcal{S}. \tag{3.101}$$

Similarly, there exists an output-nulling system with signal space $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ and set of stable trajectories $\mathcal{S}'$ such that

$$\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathcal{S}_G \iff \left[\begin{smallmatrix} u \\ -y \end{smallmatrix}\right] \in \mathcal{S}'. \tag{3.102}$$

Remark 3.6.21. The equality in (3.101) and equivalence in (3.102) are really isomorphisms, understood in a similar sense to Theorem 3.1.18.

Proof of Proposition 3.6.20: We prove the driving-variable system case first. By Varga [87], [88], G has a right coprime factorisation $G = NM^{-1}$, with M taking values in $B(\mathscr{U})$, N taking values in $B(\mathscr{U}, \mathscr{Y})$ and both proper rational. By standard realisation theory, there is an input-state-output node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ with input, state and output spaces $\mathscr{U}$, $\mathscr{X}$ and $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ respectively, with transfer function $\left[\begin{smallmatrix} M \\ N \end{smallmatrix}\right]$. Set $\mathscr{V} = \mathscr{U}$ and $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ and consider the driving-variable system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\text{dv}}$. We now claim that (3.101) holds. Let $w \in \mathcal{S}$, with corresponding state $x \in C(\mathbb{R}^+; \mathscr{X})$ and driving-variable $v \in L^2_{\text{loc}}(\mathbb{R}^+; \mathscr{V})$ so that (3.11) holds. By construction there exists $u \in L^2(\mathbb{R}^+; \mathscr{U})$ and $y \in L^2(\mathbb{R}^+; \mathscr{Y})$ such that

$$w = u + y = \begin{bmatrix} u \\ y \end{bmatrix},$$

and thus

$$\begin{aligned} \dot{x} &= Ax + Bv, \\ \begin{bmatrix} u \\ y \end{bmatrix} &= Cx + Dv = \begin{bmatrix} C_{\mathscr{U}} x \\ C_{\mathscr{Y}} x \end{bmatrix} + \begin{bmatrix} D_{\mathscr{U}} v \\ D_{\mathscr{Y}} v \end{bmatrix}. \end{aligned} \tag{3.103}$$

Taking the Laplace transform of (3.103) (using Remark 3.6.19) we obtain

$$\begin{aligned} \hat{u}(s) &= (D_{\mathscr{U}} + C_{\mathscr{U}}(sI - A)^{-1}B)\hat{v}(s) = M(s)\hat{v}(s), \\ \hat{y}(s) &= (D_{\mathscr{Y}} + C_{\mathscr{Y}}(sI - A)^{-1}B)\hat{v}(s) = N(s)\hat{v}(s), \end{aligned} \quad \forall\, s \in \mathbb{C}_\alpha^+, \tag{3.104}$$

for some $\alpha > 0$. Using the fact that M is pointwise invertible we recover

$$\begin{aligned} \hat{y}(s) &= (D_{\mathscr{Y}} + C_{\mathscr{Y}}(sI - A)^{-1}B)(D_{\mathscr{U}} + C_{\mathscr{U}}(sI - A)^{-1}B)^{-1}\hat{u}(s) \\ &= N(s)M^{-1}(s)\hat{u}(s) = G(s)\hat{u}(s), \quad \forall\, s \in \mathbb{C}_\alpha^+, \end{aligned} \tag{3.105}$$

that is, $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathcal{S}_G$. The converse inclusion essentially reverses these steps. Suppose $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathcal{S}_G$ so that (3.105) holds. Define $\hat{v}(s)$ by

$$\hat{v}(s) := M^{-1}(s)\hat{u}(s), \quad \forall\, s \in \mathbb{C}_\alpha^+,$$

so that (3.104) holds. We claim that $v(s) \in L^2_\alpha(\mathbb{R}^+; \mathscr{V})$. To see this, we note that since M and N are right coprime, [20, Lemma A.7.34] there exists $X, Y \in H^\infty_\alpha$ such that M, N, X and Y satisfy the Bezout identity

$$X(s)M(s) + Y(s)N(s) = I, \quad \forall\, s \in \mathbb{C}_\alpha^+.$$

Therefore

$$\begin{aligned} \hat{v}(s) &= [X(s)M(s) + Y(s)N(s)]\hat{v}(s) = \begin{bmatrix} X(s) & Y(s) \end{bmatrix} \begin{bmatrix} M(s)\hat{v}(s) \\ N(s)\hat{v}(s) \end{bmatrix} \\ &= \begin{bmatrix} X(s) & Y(s) \end{bmatrix} \begin{bmatrix} \hat{u}(s) \\ \hat{y}(s) \end{bmatrix}, \quad \forall\, s \in \mathbb{C}_\alpha^+. \end{aligned}$$

Since $u \in L^2_\alpha(\mathbb{R}^+; \mathscr{U}), y \in L^2_\alpha(\mathbb{R}^+; \mathscr{Y})$ and $\left[\begin{smallmatrix} X & Y \end{smallmatrix}\right]$ induces a bounded operator on L^2_α we deduce that $v \in L^2_\alpha(\mathbb{R}^+; \mathscr{V})$. Therefore from (3.104) we recover (3.103) for some continuous state x and thus $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathcal{S}$.

For the output-nulling case, again by Varga, G has a left coprime factorisation $G = K^{-1}L$, with K taking values in $B(\mathscr{Y})$, L taking values in $B(\mathscr{U}, \mathscr{Y})$ and both proper rational. By standard realisation theory, there is an input-state-output node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ with input, state and output spaces $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$, $\mathscr{X}$ and $\mathscr{Y}$ respectively, with transfer function $\left[\begin{smallmatrix} L & K \end{smallmatrix}\right]$. Set $\mathscr{V} = \mathscr{Y}$ and $\mathscr{W} = \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$ and consider the output-nulling system $(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right], \mathcal{T})_{\mathrm{on}}$. The equivalence of trajectories in (3.102) is proven similarly to the driving-variable case. □

Remark 3.6.22. 1. Given a driving-variable system, we say that a direct sum decomposition of the signal space is weakly admissible if with respect to this decomposition the stable signals of the driving-variable system are stable trajectories of an input-output system in the sense of Definition 3.6.16. A similar notion of weak admissibity is available for output-nulling systems. We do not define these notions precisely as they are not required.

2. It follows from Dai [21, Theorem 2-6.3] that rational functions are precisely the transfer functions of descriptor systems. Proposition 3.6.20 above shows that the stable input-output signals of a descriptor system are also the stable signals of a (finite-dimensional) driving-variable or output-nulling system.

Finally, we seek to derive a gap metric error bound for positive real balanced truncation for improper positive real functions. We restrict our attention to the positive real case, as bounded real functions necessarily belong to $H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ and if rational then must be proper.

Since strongly admissible pairs give rise to input-state-output systems, if we wish to consider improper rational positive real functions J in the framework of driving-variable systems then we need to broaden our view to weakly admissible pairs. We would like to define the positive real singular values for such functions, which, if formulated in terms of the optimal cost operator of an optimal control problem, will require some state-space realisation of J. As we have already said, we cannot consider input-state-output systems and instead we choose to use driving-variable systems. Similar results in this direction have been established by Reis & Stykel [70], where descriptor systems are used instead (see Remark 3.6.22).

Let J denote an improper positive real rational function. It follows (see Lemma 7.1.7) that $I + J$ is invertible and therefore $M := (I + J)^{-1}$ is proper rational. Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal input-state-output node realising M. Note that

$$J = (I + J - I) = (I - (I + J)^{-1})(I + J) = (I - M)M^{-1} =: NM^{-1},$$

is a right coprime factorisation of J, where $N := I - M$. Therefore $\left[\begin{smallmatrix} A & B \\ -C & I-D \end{smallmatrix}\right]$ is a minimal input-state-output realisation of N. Arguing as in the proof of Proposition 3.6.20 we see that the stable input-output signals of J are precisely the same as the stable signals of

$$\Sigma = \left(\left[\begin{array}{c|c} A & B \\ \hline C & D \\ -C & I - D \end{array} \right], \mathcal{T} \right)_{\mathrm{dv}},$$

(in the sense of (3.101)). Observe that the operator $\left[\begin{smallmatrix} D \\ I-D \end{smallmatrix}\right]$ is always injective and hence Σ is a driving-variable system as in Definition 3.1.6.

The signal space of Σ is $\mathscr{W} := \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{U} \end{smallmatrix}\right]$ which we equip with the indefinite inner-product

$$[\cdot, \cdot] : \mathscr{W} \times \mathscr{W} \to \mathbb{C}, \quad \left[\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right] := \left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle_{\mathscr{U} \times \mathscr{U}}. \tag{3.106}$$

As J is positive real, it follows from $\mathcal{S}_J = \mathcal{S}$ that Σ is signal dissipative (cf. Proposition 3.3.7).

We can repeat the above construction for the dual positive real transfer function J_d. Proposition 3.4.8 can be extended to weakly admissible decompositions in the obvious way, so that the following commuting diagram holds

$$\begin{array}{ccc} \left(\left[\left[\begin{smallmatrix} A \\ C \\ -C \end{smallmatrix}\right] \left[\begin{smallmatrix} B \\ D \\ I-D \end{smallmatrix}\right]\right], \mathcal{S}\right)_{\mathrm{dv}} & \xrightarrow{\text{dual}} & \left(\begin{bmatrix} A^* & -\left[\begin{smallmatrix} C \\ -C \end{smallmatrix}\right]^\dagger \\ -B^* & \left[\begin{smallmatrix} D \\ I-D \end{smallmatrix}\right]^\dagger \end{bmatrix}, \mathcal{S}^*\right)_{\mathrm{on}} \\ \text{derived}\Big\downarrow & & \Big\downarrow\text{derived} \\ (J, \mathcal{S}_J) & \xrightarrow{\text{dual}} & (J_d, \mathcal{S}_J^*), \end{array} \tag{3.107a}$$

where the trajectories are related as follows

$$\begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{S}_J \iff \begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{S} \quad \text{and} \quad \begin{bmatrix} y_* \\ u_* \end{bmatrix} \in \mathcal{S}_J^* \iff \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} -u_* \\ y_* \end{bmatrix} = \begin{bmatrix} -y_* \\ u_* \end{bmatrix} \in \mathcal{S}^*. \tag{3.107b}$$

(again, as in Remark 3.1.19 the above sets of trajectories are really isomorphic). In particular, since J_d is positive real it follows that Σ^* is signal dissipative and hence Σ is jointly signal-dissipative. We are now in position to formulate the definition of positive real singular values in the improper case.

Definition 3.6.23. Let $J : \mathbb{C}_0^+ \to B(\mathscr{U})$ denote a positive real transfer function and let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ denote a minimal input-state-output realisation of $(I+J)^{-1}$ with input, state and output spaces $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{U}$ respectively. Let Σ denote the driving-variable system

$$\left(\left[\begin{array}{c|c} A & B \\ \hline C & D \\ -C & I-D \end{array}\right], \mathcal{T}\right)_{\mathrm{dv}}. \tag{3.108}$$

(with driving-variable, state and signal spaces $\mathscr{U}$, $\mathscr{X}$ and $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{U} \end{smallmatrix}\right]$ respectively), which is jointly dissipative with respect to $[\cdot, \cdot]$ as in (3.106). We define the positive real singular values of J as the dissipative characteristic values of Σ. Let

$$\left(\left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \\ -C_1 & I-D \end{array}\right], \mathcal{T}\right)_{\mathrm{dv}}. \tag{3.109}$$

denote the dissipative balanced truncation of order r of Σ, with set of stable trajectories $\mathcal{S}_r$. If J_r given by

$$J_r(s) = (I - D - C_1(sI - A_{11})^{-1}B_1)(D + C_1(sI - A_{11})^{-1}B_1)^{-1}$$

is well-defined and satisfies $\mathcal{S}_{J_r} = \mathcal{S}_r$ then we call J_r the positive real balanced truncation.

Remark 3.6.24. Using the notation of Definition 3.6.3 we remark that if $(\mathscr{U}, \mathscr{U})$ is strongly admissible for Σ then it is strongly admissible for Σ_r as Σ and Σ_r have the same driving-variable to signal (i.e. "D") operator. We are not sure whether this is true in the so-called weakly admissible case. Namely we are not sure whether invertibility of

$$D + C(sI - A)^{-1}B,$$

implies that

$$D + C_1(sI - A_{11})^{-1}B_1,$$

is also invertible. As such we are not sure at present whether it is always possible to find the positive real balanced truncation of an improper positive real transfer function.

In the case when the positive real balanced truncation exists we obtain the same gap metric error bound as in Corollary 3.6.9.

Theorem 3.6.25. *Using the notation of Definition 3.6.23, assume that for $r < m$ the positive real balanced truncation J_r exists. Then*

$$\hat{\delta}(J, J_r) \leq 2 \sum_{k=r+1}^{m} \sigma_k,$$

where $(\sigma_i)_{i=1}^m$ are the positive real singular values of J, as in Definition 3.6.23.

Proof. Let Σ denote the jointly dissipative driving-variable system as in (3.108), with set of stable externally generated trajectories $\mathcal{S}$. By definition the positive real singular values of J are precisely the dissipative characteristic values of Σ. From Theorem 3.6.8 the gap metric error bound (3.85) holds. By the commuting diagram (3.107) it follows that $\mathcal{S}_J = \mathcal{S}$ and similarly for the dissipative balanced truncations $\mathcal{S}_{J_r} = \mathcal{S}_r$. Therefore

$$\hat{\delta}(J, J_r) = \hat{\delta}(\mathcal{S}_J, \mathcal{S}_{J_r}) = \hat{\delta}(\mathcal{S}, \mathcal{S}_r) = \hat{\delta}(\Sigma, \Sigma_r) \leq 2 \sum_{i=r+1}^{m} \sigma_i,$$

completing the proof. □

3.7 Notes

State space systems where no distinction is made between inputs and outputs are of course not new and have been studied in, for example, [4]-[8],[44], [45], [67] and [100]-[102]. Model reduction for such systems has also been considered in the literature. Particularly by Weiland [91] where LQG balancing in a behavioral framework was studied. Dissipativity retaining balanced truncation in a behavioral framework has recently been addressed by Minh [51] and Trentelman [86], [85]. However, the error bounds that are provided are, as in the input-state-output framework, on the H^∞-norm of the difference of the original and the reduced transfer function. Transfer functions are however non-behavioral objects since they depend on the input/output decomposition. The main result of this chapter, Theorem 3.6.8, is new. A version of this chapter has been submitted for publication as [38].

Model reduction for positive real and bounded real descriptor systems has been addressed in [70], where H^∞ error bounds are derived. Recall from Remark 3.6.22 that descriptor systems are precisely the realisations of rational transfer functions. We make a connection between rational transfer functions and weakly derived input-output systems in Proposition 3.6.20. We admit that our presentation of weak admissibility and model reduction of improper transfer functions is not as complete as we would have hoped, as we would like sufficient conditions for knowing that the positive real balanced truncation always exists. Accordingly, Section 3.6.3 was not included in [38].

Model reduction of bounded real and positive real input-state-output systems by singular perturbation approximation was considered by Muscato *et al.* [53]. There they show that the singular perturbation approximation retains the respective dissipativity property, and that the direct truncation error bounds translate across to the singular perturbation approximation. In Theorem 3.6.14 we demonstrate that singular perturbation approximation is often suitable in our framework and as a consequence we obtain the same gap metric error bound as for dissipative balanced truncation.

Part II

Infinite-dimensional theory

Chapter 4

Preliminaries

Here we gather together some of the material we will require for the second part of this thesis.

4.1 Well-posed linear systems

We briefly recap well-posed linear systems and realisations. Well-posed linear systems on L^2 go back to the work of Salamon [72], [73]. The more recent monograph of Staffans [81] is dedicated to the study of well-posed linear systems, and we will make use of many results from this text. Here we collect the notation we use and a few key results for well-posed linear systems. We remark that there are several different but equivalent formulations in the literature of a well-posed linear system. Although we use many results from [97], we have chosen to use the formulation of [81] so as to more readily apply results from that book. The equivalence between the formulations in [97] and [81] is shown in [81, Section 2.8].

For precise definitions of the following objects we refer the reader to [81, Section 2.2]. We denote by $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ on $(\mathscr{Y}, \mathscr{X}, \mathscr{U})$ (respectively, the output, state and input spaces) an L^p well-posed linear system with state x and output y given by

$$\begin{aligned} x(t) &= \mathfrak{A}^t x_0 + \mathfrak{B}_0^t u, \\ y &= \mathfrak{C}_0 x_0 + \mathfrak{D}_0 u, \quad t \geq 0, \\ x(0) &= x_0, \end{aligned} \tag{4.1}$$

for input $u \in L^p_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{U})$. We will mostly be using L^2 well-posed systems, though we will also need L^1 well-posed systems. In the above $(\mathfrak{A}^t)_{t\geq 0}$ is a strongly continuous semigroup on the state-space $\mathscr{X}$, $\mathfrak{B}_0^t$ is the input map (with initial time 0 and final time t), $\mathfrak{C}_0$ the output map and $\mathfrak{D}_0$ the input-output map (both with initial time 0). We remark that the finite-dimensional input-state-output system (2.3) has operators

$\mathfrak{A}^t, \mathfrak{B}_0^t, \mathfrak{C}_0$ and $\mathfrak{D}_0$ given by

$$\begin{aligned} \mathfrak{A}^t =& \mathrm{e}^{At}, & \mathfrak{B}_0^t u =& \int_0^t \mathrm{e}^{A(t-s)} Bu(s)\, ds, \\ (\mathfrak{C}_0 x_0)(t) =& C\mathrm{e}^{At} x_0, & (\mathfrak{D}_0 u)(t) =& Du(t) + C \int_0^t \mathrm{e}^{A(t-s)} Bu(s)\, ds. \end{aligned} \tag{4.2}$$

Remark 4.1.1. As explained in [81, Definition 2.2.6] and [81, Theorem 2.2.14], the operators $\mathfrak{B}_0^t$, $\mathfrak{C}_0$ and $\mathfrak{D}_0$ can be expressed in terms of the master operators $\mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$ and vice versa. There is no issue, therefore, with using the master operators $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$. For example, for the finite-dimensional system (2.3), $\mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$ are given by

$$\begin{aligned} \mathfrak{B}u &= \int_{-\infty}^0 \mathrm{e}^{-As} Bu(s)\, ds, \qquad \mathfrak{C}x = (\mathbb{R}^+ \ni t \mapsto C\mathrm{e}^{At} x), \\ \mathfrak{D}u &= \left(\mathbb{R} \ni t \mapsto \int_{-\infty}^t C\mathrm{e}^{A(t-s)} u(s)\, ds + Du(t) \right). \end{aligned}$$

Remark 4.1.2. We assume that the reader is familiar with the generators of a well-posed linear system. The control operator and observation operator of well-posed linear systems date back to Weiss, [92] and [93] respectively. We shall also require the notion of a regular transfer function, as introduced by Weiss [95], and an operator node, system node, a compatible operator node and an admissible feedback operator. All of these concepts can be found in [81] and the latter are only drawn upon in some of the proofs of our later results, and are not needed for understanding the statements of those results.

Definition 4.1.3. For $1 \leq p < \infty$ and a Banach space $\mathscr{X}$ we let π_+ and π_- denote the projections from $L^p(\mathbb{R}; \mathscr{X})$ onto $L^p(\mathbb{R}^+; \mathscr{X})$ and $L^p(\mathbb{R}^-; \mathscr{X})$ respectively (which are defined as linear operators induced by multiplication with the characteristic functions on $\mathbb{R}^+$ and $\mathbb{R}^-$ respectively). Therefore, for $u \in L^p(\mathbb{R}; \mathscr{X})$

$$\pi_+ u(t) = \begin{cases} u(t), & t \geq 0 \\ 0, & t < 0 \end{cases}, \quad \pi_- u(t) = \begin{cases} u(t), & t \leq 0 \\ 0, & t > 0. \end{cases}$$

We let τ^t denote both the bilateral shift by t on $L^p(\mathbb{R}; \mathscr{X})$, i.e. for $t, s \in \mathbb{R}$, $(\tau^t u)(s) = u(t+s)$ and the left shift by t on $L^p(\mathbb{R}^+; \mathscr{X})$, i.e. for $t, s \in \mathbb{R}^+$, $(\tau^t u)(s) = u(t+s)$. Which shift is being used will be made clear from the context. When $p = 2$ the adjoint of the left shift by $t \geq 0$ on $\mathbb{R}^+$, τ^t, is the right shift which we denote by $(\tau^t)^*$ and satisfies

$$((\tau^t)^* u)(s) = \begin{cases} u(s-t), & s \geq t \\ 0, & s < t \end{cases} \quad s \geq 0.$$

We let R denote the reflection in time (about zero), $(Rv)(t) = v(-t)$. The reflection R

acts on $L^2(\mathbb{R};\mathscr{Z})$, and we view elements of $L^2(\mathbb{R}^+;\mathscr{Z})$ or $L^2(\mathbb{R}^-;\mathscr{Z})$ as belonging to $L^2(\mathbb{R};\mathscr{Z})$ by extending by zero.

Remark 4.1.4. We adopt the convention of Hankel operators mapping forwards time to forwards time. Therefore it is necessary to include a reflection operator R (as in Definition 4.1.3) in our definition of Hankel operator of a well-posed linear system when compared to [81]. Specifically, the Hankel operator induced by a linear, time-invariant causal operator $\mathfrak{D}$ is the map $\pi_+\mathfrak{D}\pi_-R$.

The term realisation of an input-output (linear, time-invariant, causal) map $\mathfrak{D}$ on L^p refers to an L^p well-posed linear system with input-output map $\mathfrak{D}$, see [81, Definition 2.6.3] for more details. The transfer function G of an L^p well-posed system is defined as (see [81, Definition 4.6.1]) the analytic $B(\mathscr{U},\mathscr{Y})$ valued function

$$s \mapsto \big(u \mapsto \mathfrak{D}(\mathrm{e}^{st}u)(0)\big), \quad u \in \mathscr{U}, \tag{4.3}$$

defined for $\operatorname{Re} s > \omega_{\mathfrak{A}}$ (the growth bound of $\mathfrak{A}$). The transfer function G is usually understood, however, as the "Laplace transform of the input-output map," which by [81, Corollary 4.6.10] is equivalent to the above definition. We refer the reader to [81, Corollary 4.6] or Weiss [94] for more information.

The transfer function in (4.3) determines $\mathfrak{D}$ uniquely and hence by a realisation of a transfer function we mean a realisation of the input-output map $\mathfrak{D}$ related to G by (4.3).

The following result is well-known and simply states that every H^∞ function has a (stable) L^2 well-posed realisation, with Hilbert space state-space.

Lemma 4.1.5. *Given $G \in H^\infty(\mathbb{C}_0^+;B(\mathscr{U},\mathscr{Y}))$, there exists an L^2 well-posed realisation $\Sigma = (\mathfrak{A},\mathfrak{B},\mathfrak{C},\mathfrak{D})$ on $(\mathscr{Y},\mathscr{X},\mathscr{Y})$ with $\mathscr{X}$ a Hilbert space such that the following stability assumptions hold:*

$$\mathfrak{A},\mathfrak{A}^* \;\; \textit{are strongly stable}, \tag{4.4}$$

$$\mathfrak{B} : L^2(\mathbb{R}^-;\mathscr{U}) \to \mathscr{X} \;\; \textit{is bounded}, \tag{4.5}$$

$$\mathfrak{C} : \mathscr{X} \to L^2(\mathbb{R}^+;\mathscr{Y}) \quad \textit{is bounded}, \tag{4.6}$$

$$\mathfrak{D} : L^2(\mathbb{R};\mathscr{U}) \to L^2(\mathbb{R};\mathscr{Y}) \quad \textit{is bounded}. \tag{4.7}$$

As such the operator $\mathfrak{D}_0 = \mathfrak{D}\pi_+ : L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y})$ is also bounded. We call such a system (in particular satisfying (4.4)-(4.7)*) a stable L^2 well-posed linear system.*

Proof. This is well-known and follows from, for example, [97, Theorem 4.2]. □

4.1.1 Shift realisations

We collect an important family of shift realisations that we will make frequent use of in our balanced truncation work.

Lemma 4.1.6. *For a linear, time-invariant, causal operator* $\mathfrak{D} : L^p(\mathbb{R}; \mathscr{U}) \to L^p(\mathbb{R}; \mathscr{Y})$ *with* $1 \leq p < \infty$ *and* $\mathscr{U}$, $\mathscr{Y}$ *Banach spaces the system*

$$^{sr}\Sigma^p = (\tau, HR, I, \mathfrak{D}), \quad \text{on } (\mathscr{Y}, L^p(\mathbb{R}^+; \mathscr{Y}), \mathscr{U}),$$

is an L^p *well-posed linear system. Here* τ *and* I *are the left-shift and identity on* $L^p(\mathbb{R}^+; \mathscr{Y})$ *respectively, and* $H = \pi_+ \mathfrak{D} \pi_- R$ *is the Hankel operator. We call* $^{sr}\Sigma^1$ *the exactly observable shift realisation on* L^1 *of* $\mathfrak{D}$ *and* $^{sr}\Sigma^2$ *the output-normal shift realisation on* L^2 *of* $\mathfrak{D}$.

Proof. That $^{sr}\Sigma^p$ is an L^p well-posed linear system follows from [81, Example 2.6.5 (ii)] (noting our convention in Remark 4.1.4 for Hankel operators). Note that the left-shift τ is a strongly continuous semigroup on $L^p(\mathbb{R}^+; \mathscr{Y})$ by [81, Example 2.3.2 (ii)]. □

Lemma 4.1.7. *Let* τ *denote the strongly continuous left-shift semigroup on* $L^p(\mathbb{R}^+; \mathscr{X})$, *where* $1 \leq p < \infty$ *and* $\mathscr{X}$ *is a Banach space. The generator of* τ *is*

$$A : D(A) \to L^p(\mathbb{R}^+; \mathscr{X}), \quad A = \frac{d}{dt}, \quad D(A) = W^{1,p}(\mathbb{R}^+; \mathscr{X}),$$

where $W^{1,p}$ *is the usual Sobolev space. When* $p = 2$ *the adjoint of* τ, *the right-shift operator* τ^* *is also a strongly continuous semigroup and has generator*

$$A^* : D(A^*) \to L^2(\mathbb{R}^+; \mathscr{X}), \quad A^* = -\frac{d}{dt}, \quad D(A^*) = W_0^{1,2}(\mathbb{R}^+; \mathscr{X}),$$

where $W_0^{1,2}(\mathbb{R}^+; \mathscr{X}) = \{w \in W^{1,2}(\mathbb{R}^+; \mathscr{X}) \; : \; w(0) = 0\}$.

Proof. See [81, Example 3.2.3 (ii)] for the generator of the left shift and the generator of the right shift can be inferred from [81, Example 3.2.3 (iii)]. □

4.2 Dual systems

For bounded real and positive real balanced truncation given an H^∞ transfer function with L^2 well-posed realisation we will need the following notion of a dual transfer function and dual well-posed realisation.

Definition 4.2.1. Given Hilbert spaces $\mathscr{U}$ and $\mathscr{Y}$ and transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, the dual transfer function $G_d \in H^\infty(\mathbb{C}_0^+; B(\mathscr{Y}, \mathscr{U}))$ is defined as

$$\mathbb{C}_0^+ \ni s \mapsto G_d(s) = [G(\overline{s})]^*. \tag{4.8}$$

Given an L^2 well-posed linear system $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ on $(\mathscr{Y}, \mathscr{X}, \mathscr{U})$ (Hilbert spaces) we call the L^2 well-posed linear system Σ_d given by

$$\Sigma_d = ({}^d\mathfrak{A}, {}^d\mathfrak{B}, {}^d\mathfrak{C}, {}^d\mathfrak{D}) = (\mathfrak{A}^*, \mathfrak{C}^* R, R\mathfrak{B}^*, R\mathfrak{D}^* R),$$

on $(\mathscr{U}, \mathscr{X}, \mathscr{Y})$ the (causal) dual of Σ. Here R is the reflection in time from Definition 4.1.3. Given an input $y_d \in L^2_{\mathrm{loc}}(\mathbb{R}^+; \mathscr{Y})$, the state x_d and output u_d of Σ_d are defined by

$$\begin{aligned} x_d(t) &= (\mathfrak{A}^t)^* x_0 + {}^d\mathfrak{B}_0^t y_d, \\ u_d &= {}^d\mathfrak{C} x_0 + {}^d\mathfrak{D}_0 y_d, \qquad t \geq 0, \\ x_d(0) &= x_0. \end{aligned} \tag{4.9}$$

Remark 4.2.2. In [81, Theorem 6.2.3] it is proven that Σ_d is an L^2 well-posed linear system .

The following result describes some properties of dual systems, notably that the dual system realises the dual transfer function, and is again taken from [81].

Lemma 4.2.3. *Let $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ denote an L^2 well-posed linear system on $(\mathscr{Y}, \mathscr{X}, \mathscr{U})$ and let (A, B, C) and G denote the generators and transfer function of Σ respectively. The dual system Σ_d has generators (A^*, C^*, B^*) and transfer function G_d. If Σ is stable, then so is Σ_d.*

Proof. The claims regarding the generators and transfer function of Σ_d follow from [81, Theorem 6.2.13]. That Σ_d is stable follows from the definition of the operators ${}^d\mathfrak{A}$, ${}^d\mathfrak{B}$, ${}^d\mathfrak{C}$ and ${}^d\mathfrak{D}$ and the stability of Σ. $\square$

4.3 Singular values, nuclear operators and Schmidt pairs

Definition 4.3.1. For a bounded linear operator $T : \mathscr{Z}_1 \to \mathscr{Z}_2$ between Banach spaces, the i^{th} singular value s_i (also called s-value or approximation number) is defined as

$$s_i := \inf \{ \|T - S\| \ : \ \operatorname{rank} S \leq i - 1 \} .$$

The operator T is nuclear if its singular values $(s_i)_{i \in \mathbb{N}}$ are summable, that is

$$\sum_{i \in \mathbb{N}} s_i < \infty. \tag{4.10}$$

For a nuclear operator T, we call the series in (4.10) the nuclear norm of T, which we denote by $\|T\|_N$.

If $T : \mathscr{Z}_1 \to \mathscr{Z}_2$ is nuclear with singular values $(s_i)_{i \in \mathbb{N}}$ then necessarily

$$s_i \to 0, \quad \text{as} \quad i \to \infty.$$

In particular for each $i \in \mathbb{N}$ there exists a bounded operator $S(i)$ of rank $\leq i-1$ such that

$$\|T - S(i)\| \leq s_i + \frac{1}{i} \to 0, \quad \text{as} \quad i \to \infty.$$

We see that a nuclear operator T is the uniform limit of finite rank operators and hence is compact.

Definition 4.3.2. For $\mathscr{X}_1, \mathscr{X}_2$ Banach spaces the Schatten class $S_p(\mathscr{X}_1, \mathscr{X}_2)$ is the set of operators whose singular values form an ℓ^p sequence. In particular, $S_1(\mathscr{X}_1, \mathscr{X}_2)$ is the class of nuclear operators.

Remark 4.3.3. Some authors use the term trace instead of nuclear, and define nuclear operators as compact operators with summable singular values, possibly also defining the singular values as the positive square roots of the eigenvalues of T^*T when $\mathscr{X}_1, \mathscr{X}_2$ are Hilbert spaces, for example, Lax [47, p. 330]. This definition of singular value is equivalent to that above by Gohberg *et al.* [29, Theorem VI. 1.5].

Remark 4.3.4. In this work we use the term singular value in a non-standard manner. Given Definition 4.3.1 above, we call σ_i the i^{th} singular value of T, but counted with multiplicities, so that if $s_1 = s_2 = \cdots = s_{p_1}$ and $s_{p_1} > s_{p_1+1}$, for some $p_1 \in \mathbb{N}$ then we set

$$\sigma_1 = s_1 = s_2 = \cdots = s_{p_1}, \quad \sigma_2 = s_{p_1+1} = \ldots,$$

and so on. As such, our i^{th} singular value σ_i has multiplicity $p_i \in \mathbb{N}$ and satisfies $\sigma_i > \sigma_{i+1}$, however note that σ_i need not necessarily be the distance of T from rank $i-1$ operators. Using this convention the operator T is nuclear if

$$\sum_{i \in \mathbb{N}} p_i \sigma_i < \infty.$$

We remark that if all the singular values are simple (i.e. all have multiplicity one), then our convention and the usual convention coincide.

Definition 4.3.5. For a compact linear operator $T : \mathscr{X}_1 \to \mathscr{X}_2$ between Hilbert spaces, let $(\sigma_k)_{k \in \mathbb{N}}$ denote the singular values of T (which are precisely the non-negative square roots of the countably many eigenvalues of T^*T), each with multiplicity $p_i \in \mathbb{N}$. The Schmidt pairs $(v_{i,k}, w_{i,k})$ of T are eigenvectors of T^*T and TT^* respectively corresponding to the eigenvalue σ_i^2. The Schmidt pairs can be chosen to satisfy

$$\left.\begin{aligned} w_{i,k} \in \mathscr{X}_2, \quad Tv_{i,k} &= \sigma_i w_{i,k}, \\ v_{i,k} \in \mathscr{X}_1, \quad T^* w_{i,k} &= \sigma_i v_{i,k}, \end{aligned}\right\} \quad \forall\, i \in \mathbb{N},\ 1 \leq k \leq p_i,$$

and are always chosen orthonormal.

4.4 Notes

Well-posed linear systems as presented here are based on the work of Salamon [72], [73], and have subsequently been contributed to by many authors. We refer the reader to [81, Section 2.9] for a more thorough history of the development of the field. As Staffans in [81, p.78] says, however, "the class of L^p well-posed linear systems which we present ... is by no means the only possible setting for an infinite-dimensional systems theory." More general infinite-dimensional systems include system and operator nodes, introduced by Šmuljan [75], and distributional linear systems as introduced by Opmeer [59]. The class of L^p well-posed linear systems is sufficiently general for our needs and (as Lemma 4.1.5 demonstrates) a suitable environment for realising H^∞ transfer functions. Since, broadly speaking, in Part II of this thesis we are looking to develop model reduction by balanced truncation for transfer functions belonging to H^∞, we stick to this framework. In fact, key results from the literature that we use in deriving bounded real and positive real balanced truncation are presented for well-posed linear systems.

The material in Section 4.3 of this chapter is of course well-known and can be found in many textbooks. We particularly used Gohberg *et al.* [29], Lax [47] and Swartz [83].

Chapter 5

Lyapunov balanced truncation

In this chapter we describe some extensions of Lyapunov balanced truncation, described for finite-dimensional systems in Section 2.1, to the infinite-dimensional case. Specifically, for $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, no longer necessarily rational, we seek a rational $G_n \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ such that

$$\|G - G_n\|_{H^\infty} \leq 2 \sum_{k \geq n+1} \sigma_k, \tag{5.1}$$

where σ_k are the Lyapunov singular values[1]. Here $\mathscr{U}$ and $\mathscr{Y}$ are finite-dimensional Hilbert spaces. In Glover *et al.* [27] balanced truncations and the H^∞ error bound (5.1) were extended to a class of infinite-dimensional systems.

The error bound (5.1) obviously holds when the right-hand side is infinite, and so there is only something to prove for systems such that $\sum_{k \geq n+1} \sigma_k < +\infty$ (for some or equivalently for all n), i.e. systems that have a nuclear Hankel operator. In proving the error bound (5.1), the authors of [27] made further assumptions that are unnecessarily restrictive. Specifically, their assumptions are

(i) The Hankel operator $H : L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$ is of the form

$$(Hf)(t) = \int_{\mathbb{R}^+} h(t+s) f(s)\, ds, \quad f \in L^2(\mathbb{R}^+; \mathscr{U}), \quad a.a.\ t \geq 0, \tag{5.2}$$

with kernel $h \in L^1(\mathbb{R}^+; B(\mathscr{U}, \mathscr{Y}))$.

(ii) $h \in L^2(\mathbb{R}^+; B(\mathscr{U}, \mathscr{Y}))$.

(iii) The kernel h is real, or if h is complex then $\dot{h}$ exists and $\dot{h}$ is the kernel of a bounded Hankel operator.

(iv) The singular values of the Hankel operator are assumed to be simple.

[1] That is, the singular values of the Hankel operator, see Remark 2.1.7.

(v) The Hankel operator H is also assumed to be nuclear.

We give an example of a physical system where assumption (ii) fails.

Example 5.0.1. Consider the following heat equation in 1D on the unit interval

$$w_t(t,x) = w_{xx}(t,x), \quad t \geq 0,\ x \in [0,1]. \tag{5.3a}$$

We impose Neumann control and Dirichlet observation at the left end, i.e. for $t \geq 0$

$$u(t) := -w_x(t,0), \tag{5.3b}$$

$$y(t) := w(t,0), \tag{5.3c}$$

(so that the input and output spaces are one-dimensional) and a Dirichlet boundary condition at the right end

$$w(t,1) = 0, \quad \forall\, t \geq 0. \tag{5.3d}$$

It follows as in Opmeer [61, Section 3, Theorem 3] that the system (5.3) has Schatten class S_p Hankel operator for all $p > 0$ and thus in particular the Hankel operator is nuclear.

To find the transfer function G we take the Laplace transform of (5.3a) and solve the resulting ODE (with zero initial temperature profile $w(0,x) = 0,\ \forall x \in [0,1]$), which can be justified by Curtain & Zwart [20, Examples 4.3.11, 4.3.12]. Some elementary calculations give

$$J(s) = \frac{1}{\sqrt{s}} \tanh(\sqrt{s}), \quad \operatorname{Re} s > 0, \tag{5.4}$$

where we take the (unique) square root $\sqrt{s}$ with argument in $(-\frac{\pi}{2}, \frac{\pi}{2})$, so that $\operatorname{Re}\sqrt{s} > 0$.

We claim that $J \notin H^2(\mathbb{C}_0^+)$ and hence the impulse response $h \notin L^2(\mathbb{R}^+)$. To see this, note that along the line $\{t(1+\mathrm{i}) \,:\, t \geq 0\}$ we have

$$|\tanh(t(1+\mathrm{i}))|^2 \to 1 \quad \text{as } t \to \infty. \tag{5.5}$$

Therefore

$$\begin{aligned}
\|J\|_{H^2}^2 &= \int_{\mathbb{R}} \left| \frac{1}{\sqrt{\omega \mathrm{i}}} \tanh(\sqrt{\omega \mathrm{i}}) \right|^2 d\omega = \int_{\mathbb{R}} \frac{1}{|\omega|} \left|\tanh(\sqrt{\omega \mathrm{i}})\right|^2 d\omega, \\
&\geq \int_{\mathbb{R}^+} \frac{1}{\omega} \left|\tanh\left(\sqrt{\omega}(\tfrac{1}{\sqrt{2}} + \tfrac{\mathrm{i}}{\sqrt{2}})\right)\right|^2 d\omega, \\
&\geq \int_C^\infty \frac{1}{2\omega}\, d\omega,
\end{aligned} \tag{5.6}$$

by (5.5), for some $C > 0$ sufficiently large. The integral (5.6) is infinite and hence $J \notin H^2$. It follows that the results of [27] therefore do not apply to this example.

We extend bounded real (and hence positive real) balanced truncation to the infinite-dimensional case by relating it to Lyapunov balanced truncation, as described for the finite-dimensional case in Section 2.2.3. However, the results of [27] require assumptions that are too strong (and as the above example demonstrates aren't always fulfilled in practice). Therefore, in this chapter we consider realisation and approximation of transfer functions with a nuclear Hankel operator and contrary to [27], *we impose no further restrictions.* In doing so we also carefully describe where the assumptions (i) -(v) are used in [27]. As mentioned, we use these results in deriving bounded real and positive balanced truncation in the infinite-dimensional case, but are also of independent interest.

The main results of this chapter are the following two theorems.

Theorem 5.0.2. *If H is a nuclear Hankel operator with transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, where $\mathscr{U}$ and $\mathscr{Y}$ are finite-dimensional, then for any positive integer n*

$$\|G - G_n\|_{H^\infty} \leq 2 \sum_{k=n+1}^{\infty} \sigma_k, \tag{5.7}$$

where σ_k are the singular values of H and G_n is the reduced order transfer function obtained by Lyapunov balanced truncation from Definition 5.3.5.

Proof. Theorem 5.0.2 is proven in Section 5.4. □

Definition 5.3.5 does not provide a constructive method of finding G_n, as G_n is defined as the transfer function of a realisation that depends on the Schmidt vectors of the Hankel operator H, which cannot usually be found in practice. Our second result demonstrates that for the class of systems with Hankel operators given by (5.2) (which by Corollary 5.1.18 includes nuclear Hankel operators), G_n is the H^∞ limit of a sequence of computable transfer functions.

Theorem 5.0.3. *Suppose H is a Hankel operator given by (5.2) with L^1 kernel h, transfer function G, and reduced order transfer function G_n obtained by Lyapunov balanced truncation. Suppose $(h_m)_{m\in\mathbb{N}}$ is such that*

$$h_m \xrightarrow{L^1} h, \quad as \quad m \to \infty,$$

and let $(G^m)_{n\in\mathbb{N}}$ denote the corresponding sequence of transfer functions. If $(G_n^m)_{m\in\mathbb{N}}$ denotes the sequence of reduced order transfer functions obtained from G^m by Lyapunov balanced truncation, then there exists a subsequence $(\tau(s))_{s\in\mathbb{N}}$ such that

$$G_n^{\tau(s)} \xrightarrow{H^\infty} G_n, \quad as \quad s \to \infty. \tag{5.8}$$

Proof. See Proposition 5.3.9. □

Remark 5.0.4. If the singular values of H are all simple then the conclusion of Theorem 5.0.3 can be replaced by

$$G_n^m \xrightarrow{H^\infty} G_n, \quad \text{as} \quad m \to \infty.$$

In other words, the convergence in (5.8) does not require a subsequence when all the singular values are simple.

Remark 5.0.5. We remark that the assumptions of Theorem 5.0.2 are almost optimal, but not quite. This is because the singular values of the Hankel operator *are not* repeated according to multiplicity in the error bound (5.7). According to Ober, Treil′ [55], [84] any sequence of non-negative real numbers forms the sequence of singular values of a Hankel operator. In particular, consider the Hankel operator with singular values $\sigma_k = \frac{1}{k^2}$, each with multiplicity k, for every $k \in \mathbb{N}$. This operator is not nuclear as

$$\sum_{k\in\mathbb{N}} k\frac{1}{k^2} = \sum_{k\in\mathbb{N}} \frac{1}{k} = \infty,$$

but the error bound

$$\sum_{k\in\mathbb{N}} \frac{1}{k^2},$$

is finite. We do not know what happens in such instances, and hope that such operators do not occur in physically motivated systems.

5.1 Hankel operators between $L^2(\mathbb{R}^+;\mathscr{Z})$ spaces

There is a large literature on Hankel operators, see for example [54], [62], [63] and [65], with unfortunately several different conventions. The purpose of this section is to first translate some of these known results into the convention used here and then to use those results to show that a nuclear Hankel operator $L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y})$ is necessarily an integral operator with L^1 kernel (Corollary 5.1.18).

We start by recalling the definition of a Hankel operator in an abstract Hilbert space setting based on shift operators. The next definition is taken from Rosenblum & Rovnyak [71, p. 1].

Definition 5.1.1. Let $\mathscr{H}$ denote a Hilbert space. An operator $S \in B(\mathscr{H})$ is a shift on $\mathscr{H}$ if S is an isometry and $(S^*)^n$ converges strongly to zero as n tends to infinity. We define a (bounded) S-Hankel operator H on $\mathscr{H}$ as a bounded operator which satisfies

$$S^*H = HS. \tag{5.9}$$

We can also define Hankel operators between two different spaces. If $\mathscr{L}$ is another Hilbert space, the operator $H \in B(\mathscr{H},\mathscr{L})$ is (S_1, S_2)-Hankel if there exist shift operators $S_1 \in B(\mathscr{H})$, $S_2 \in B(\mathscr{L})$ such that

$$S_2^* H = H S_1. \tag{5.10}$$

We seek to collect results for Hankel operators from $L^2(\mathbb{R}^+; \mathscr{Z}_1)$ to $L^2(\mathbb{R}^+; \mathscr{Z}_2)$, where $\mathscr{Z}_1$, $\mathscr{Z}_2$ are arbitrary Hilbert spaces, corresponding to the left-shift semigroup τ from Definition 4.1.3. For clarity we denote by τ_i the left-shift on $L^2(\mathbb{R}^+; \mathscr{Z}_i)$.

Definition 5.1.2. For $\mathscr{Z}_1$, $\mathscr{Z}_2$ Hilbert spaces the bounded operator $H : L^2(\mathbb{R}^+; \mathscr{Z}_1) \to L^2(\mathbb{R}^+; \mathscr{Z}_2)$ is (τ_1, τ_2)-Hankel if it is $((\tau_1^t)^*, (\tau_2^t)^*)$-Hankel for every $t \geq 0$, i.e.

$$\tau_2^t H = H(\tau_1^t)^*, \quad \forall\, t \geq 0.$$

We demonstrate in the next lemma how we can associate a (usually operator valued) Hankel matrix with an (τ_1, τ_2)-Hankel operator. From now on all bounded Hankel operators will be (τ_1, τ_2)-Hankel and so we shall omit the (τ_1, τ_2)-.

Lemma 5.1.3. *Let H denote a bounded Hankel operator $L^2(\mathbb{R}^+; \mathscr{Z}_1) \to L^2(\mathbb{R}^+; \mathscr{Z}_2)$ for $\mathscr{Z}_1, \mathscr{Z}_2$ Hilbert spaces and let A_i^* denote the generator of the right-shift semigroups $(\tau_i)^*$ from Lemma 4.1.7. Define the cogenerators (with parameter $\frac{1}{2}$) T_i of A_i^* by*

$$T_i := (A_i^* + \tfrac{1}{2} I)(A_i^* - \tfrac{1}{2} I)^{-1}, \tag{5.11}$$

which are well defined and bounded on $L^2(\mathbb{R}^+; \mathscr{Z}_i)$. Let $(u_j^i)_{j=1}^{\dim \mathscr{Z}_i}$ denote an orthonormal basis for $\mathscr{Z}_i$ and define for $k \in \mathbb{N}_0$

$$\begin{aligned} e_k u_j^i &:= T_i^k (e_0 u_j^i), \qquad \textit{where} \quad e_0(t) = \mathrm{e}^{-\frac{t}{2}}, \quad t \geq 0, \\ \mathcal{B}_0^i &:= (e_k u_j^i)_{j=1, k \in \mathbb{N}_0}^{\dim \mathscr{Z}_i}. \end{aligned}$$

Then $\mathcal{B}_0^1$ and $\mathcal{B}_0^2$ are orthonormal bases for $L^2(\mathbb{R}^+; \mathscr{Z}_1)$ and $L^2(\mathbb{R}^+; \mathscr{Z}_2)$ respectively. Moreover, define the bounded linear operators

$$h_{ij} = P_{e_j \mathscr{Z}_2} H|_{e_i \mathscr{Z}_1} : e_i \mathscr{Z}_1 \to e_j \mathscr{Z}_2, \quad \forall\, i, j \in \mathbb{N}_0,$$

where $P_{e_j \mathscr{Z}_2}$ denotes the orthogonal projection of $L^2(\mathbb{R}^+; \mathscr{Z}_2)$ onto $e_j \mathscr{Z}_2$. Then $h_{ij} = h_{i+j}$ and with respect to the bases $(\mathcal{B}_0^1, \mathcal{B}_0^2)$ the operator H has the Hankel matrix representation

$$H e_k u_r^1 = \sum_{n \in \mathbb{N}_0} h_{n+k}(e_n u_r^1), \quad \forall\, k \in \mathbb{N}_0,\ 1 \leq r \leq \dim \mathscr{Z}_1.$$

Proof. The cogenerators T_i are well defined and bounded as

$$T_i = -(A_i^* + \tfrac{1}{2} I) R_{A_i^*}(\tfrac{1}{2}),$$

where $R_{A_i^*}$ is the resolvent of $(\tau_i)^*$, which is defined (and bounded) at $s = \frac{1}{2}$. The remaining claims follow as in Adamjan *et al.* [1, p.64-65]. A direct calculation also shows that $T_i = S_i$, the Laguerre shift and $\mathcal{B}_0^i$ is the Laguerre basis on $L^2(\mathbb{R}^+; \mathscr{Z}_i)$ see, for example, Rosenblum & Rovnyak [71, p. 17] for details of the scalar case. The vectorial case is similar. □

Remark 5.1.4. Note that presently h_n defined in the proof of Lemma 5.1.3 are operators $e_i\mathscr{Z}_1 \to e_j\mathscr{Z}_2$, but these are isomorphic to operators $\tilde{h}_n : \mathscr{Z}_1 \to \mathscr{Z}_2$ (which we may view as matrices with respect to the bases $(u_j^i)_{j=1}^{\dim \mathscr{Z}_i}$). In what follows we will identify h_n with $\tilde{h}_n$.

The next result describes how a Hankel operator from $L^2(\mathbb{R}^+; \mathscr{Z}_1)$ to $L^2(\mathbb{R}^+; \mathscr{Z}_2)$ is unitarily equivalent to a linear operator from $H^2(\mathbb{C}_0^+; \mathscr{Z}_1)$ to $H^2(\mathbb{C}_0^+; \mathscr{Z}_2)$ induced by multiplication.

Lemma 5.1.5. *Let H denote a bounded Hankel operator $L^2(\mathbb{R}^+; \mathscr{Z}_1) \to L^2(\mathbb{R}^+; \mathscr{Z}_2)$ for $\mathscr{Z}_1, \mathscr{Z}_2$ Hilbert spaces. Then there exists a function $\phi \in L^\infty(\mathrm{i}\mathbb{R}; B(\mathscr{Z}_1, \mathscr{Z}_2))$ such that the operator*

$$H_\phi : H^2(\mathbb{C}_0^+; \mathscr{Z}_1) \to H^2(\mathbb{C}_0^+; \mathscr{Z}_2), \quad H_\phi = \mathcal{P}_+ M_\phi R_C, \tag{5.12}$$

satisfies

$$\mathcal{L} H \mathcal{L}^{-1} = H_\phi, \tag{5.13}$$

where we use $\mathcal{L}$ to denote both (unilateral) Laplace transforms $L^2(\mathbb{R}^+; \mathscr{Z}_i) \to H^2(\mathbb{C}_0^+; \mathscr{Z}_i)$. Furthermore $\mathcal{P}_+$ denotes the orthogonal projections of $L^2(\mathrm{i}\mathbb{R}; \mathscr{Z}_i)$ onto $H^2(\mathbb{C}_0^+; \mathscr{Z}_i)$, M_ϕ is the linear operator induced by multiplication with ϕ and R_C is reflection

$$(R_C\psi)(s) = \psi(-s).$$

Proof. We require the following Möbius transform M defined by

$$M(z) = \frac{\frac{1}{2}(1-z)}{1+z}, \quad \text{which maps} \quad \begin{array}{l} \mathbb{T} \to \mathrm{i}\mathbb{R}, \\ \mathbb{D} \to \mathbb{C}_0^+, \\ \mathbb{C} \setminus \overline{\mathbb{D}} \to \mathbb{C}_0^-. \end{array} \tag{5.14}$$

The inverse is given by

$$M^{-1}(s) = \frac{\frac{1}{2} - s}{\frac{1}{2} + s}, \quad \text{which maps} \quad \begin{array}{l} \mathrm{i}\mathbb{R} \to \mathbb{T}, \\ \mathbb{C}_0^+ \to \mathbb{D}, \\ \mathbb{C}_0^- \to \mathbb{C} \setminus \overline{\mathbb{D}}. \end{array} \tag{5.15}$$

We also need the operators $V_{\mathscr{Z}_1}, V_{\mathscr{Z}_2}$ mapping the Hardy spaces of the disc to the Hardy spaces of the right half-plane. These functions are given by

$$V_{\mathscr{X}_i} : H^2(\mathbb{D}; \mathscr{X}_i) \to H^2(\mathbb{C}_0^+; \mathscr{X}_i), \quad (V_{\mathscr{X}_i} f)(s) = \frac{1}{\frac{1}{2} + s} f(M^{-1}(s)), \quad s \in \mathbb{C}_0^+,$$

with inverse

$$V_{\mathscr{X}_i}^{-1} : H^2(\mathbb{C}_0^+; \mathscr{X}_i) \to H^2(\mathbb{D}; \mathscr{X}_i), \quad (V_{\mathscr{X}_i}^{-1} F)(z) = \frac{1}{1+z} F(M(z)), \quad z \in \mathbb{D}.$$

As described by Partington [63, p. 23-25], $V_{\mathscr{X}_i}$ are linear isomorphisms, which are up to a multiplicative constant isometric.

The Hankel matrix $H_M = (h_{i+j})_{ij}$ from Lemma 5.1.3, which by that Lemma represents H with respect to the basis $(\mathcal{B}_0^1, \mathcal{B}_0^2)$. By Nehari's Theorem in the vectorial case (contained in Peller [65, Theorem 2.2.1] for instance) H_M, considered as a linear operator $\ell^2(\mathbb{N}_0; \mathscr{X}_1) \to \ell^2(\mathbb{N}_0; \mathscr{X}_2)$ by multiplication, is a bounded operator if and only if there exists $g \in L^\infty(\mathbb{T}; B(\mathscr{X}_1, \mathscr{X}_2))$ such that

$$g(z) = \sum_{n \in \mathbb{Z}} a_n z^n, \quad \text{and} \quad a_n = h_n, \quad \forall\, n \in \mathbb{N}_0. \tag{5.16}$$

Recall that

$$\{z \mapsto z^n u_j^i \;:\; n \in \mathbb{Z},\ 1 \le j \le \dim \mathscr{X}_i\} \quad \text{and} \quad \{z \mapsto z^n u_j^i \;:\; n \in \mathbb{N}_0,\ 1 \le j \le \dim \mathscr{X}_i\},$$

are orthogonal bases for $L^2(\mathbb{T}; \mathscr{X}_i)$ and $H^2(\mathbb{D}; \mathscr{X}_i)$ respectively. We denote the above bases for $H^2(\mathbb{D}; \mathscr{X}_1)$ and $H^2(\mathbb{D}; \mathscr{X}_2)$ by $\mathcal{B}_1^1$ and $\mathcal{B}_1^2$ respectively. Standard techniques show that

$$\mathcal{B}_2^i := \left\{ V_{\mathscr{X}_i}(z^n u_j^i) \;:\; n \in \mathbb{N}_0,\ 1 \le j \le \dim \mathscr{X}_i, \right\}, \tag{5.17}$$

are orthogonal bases for $H^2(\mathbb{C}_0^+; \mathscr{X}_i)$ (see for example Partington [63, p. 23-25]). Let R_D denote the discrete time reflection

$$R_D z^n = z^{-n}, \quad n \in \mathbb{N}_0.$$

From [63, Theorem 3.1] and (5.17) we see that

$$\begin{aligned} &H_1 : H^2(\mathbb{D}, \mathscr{X}_1) \to H^2(\mathbb{D}, \mathscr{X}_2), && H_1 := P_+ M_g R_D, \\ &H_2 : H^2(\mathbb{C}_0^+, \mathscr{X}_1) \to H^2(\mathbb{C}_0^+, \mathscr{X}_2), && H_2 := V_{\mathscr{X}_2} H_1 V_{\mathscr{X}_1}^{-1}, \end{aligned}$$

have matrix representation H_M with respect to the bases $(\mathcal{B}_1^1, \mathcal{B}_1^2)$ and $(\mathcal{B}_2^1, \mathcal{B}_2^2)$ respectively. A calculation shows that

$$H_2 := V_{\mathscr{X}_2} H_1 V_{\mathscr{X}_1}^{-1} = V_{\mathscr{X}_2} P_+ M_g R_D V_{\mathscr{X}_1}^{-1} = \mathcal{P}_+ M_\phi R_C,$$

where

$$\begin{aligned} \phi(s) &= g(M^{-1}(s))\frac{M^{-1}(s)}{2}, \\ \text{and} \quad M_\phi &= V_{\mathscr{Z}_2} M_g M_{\frac{z}{2}} V_{\mathscr{Z}_1}^{-1} : L^2(\mathrm{i}\mathbb{R}; \mathscr{Z}_1) \to L^2(\mathrm{i}\mathbb{R}; \mathscr{Z}_2). \end{aligned} \tag{5.18}$$

Setting $H_\phi := H_2$ gives (5.12). Finally using

$$(V_{\mathscr{Z}_i} z^n u_j^i)(s) = \frac{(\frac{1}{2} - s)^n}{(\frac{1}{2} + s)^{n+1}} u_j^i, \quad \forall\, s \in \mathbb{C}_0^+,\ \forall\, j, n \in \mathbb{N}_0$$

a calculation shows that

$$e_n u_j^i = (-1)^n \mathcal{L}(V_{\mathscr{Z}_1} z^n u_j^i), \quad \forall\, j, n \in \mathbb{N}_0.$$

Therefore the operator

$$H_3 : L^2(\mathbb{R}^+; \mathscr{Z}_1) \to L^2(\mathbb{R}^+; \mathscr{Z}_2), \quad H_3 := \mathcal{L}^{-1} H_2 \mathcal{L},$$

has the matrix representation H_M with respect to the bases $(\mathcal{B}_0^1, \mathcal{B}_0^2)$ and hence $H_3 = H$ which proves (5.13). $\square$

5.1.1 Symbols and transfer functions

Definition 5.1.6. Using the notation of Lemma 5.1.5, a function ψ satisfying (5.12),

$$\mathcal{L} H \mathcal{L}^{-1} = H_\psi = \mathcal{P}_+ M_\psi R_C,$$

is called a symbol for the Hankel operator H. By Lemma 5.1.5, a bounded Hankel operator H always has a symbol $\phi \in L^\infty(\mathrm{i}\mathbb{R}; B(\mathscr{Z}_1, \mathscr{Z}_2))$.

We seek to relate a symbol of a bounded Hankel operator to its transfer function (what we mean by transfer function in this context is stated precisely in Definition 5.1.10). In order to do so we need two lemmas.

Lemma 5.1.7. *Let $\mathscr{Z}$ denote a Banach space. Any function $\phi \in L^\infty(\mathrm{i}\mathbb{R}; \mathscr{Z})$ has a decomposition*

$$\phi = \phi_1 + \phi_2,$$

where ϕ_1 can be extended analytically to the right half-plane and ϕ_2 can be extended analytically to the left-half plane. The components ϕ_1 and ϕ_2 are unique up to an additive constant. We refer to ϕ_1 as the analytic part of ϕ in $\mathbb{C}_0^+$.

Proof. Choose $\phi \in L^\infty(\mathrm{i}\mathbb{R}; \mathscr{Z})$ and define $f := \phi \circ M : \mathbb{T} \to \mathscr{Z}$, where M is the Möbius transform from the proof of Lemma 5.1.5. Then $f \in L^\infty(\mathbb{T}; \mathscr{Z})$ and as

$$L^\infty(\mathbb{T}, \mathscr{Z}) \subseteq L^2(\mathbb{T}, \mathscr{Z}),$$

it follows that f has a unique decomposition

$$\begin{aligned} f(z) &= \sum_{n\in\mathbb{N}_0} a_n z^n + \sum_{n\in\mathbb{N}} a_{-n} z^{-n}, \\ &=: f_1(z) + f_2(z), \quad z \in \mathbb{T}. \end{aligned} \tag{5.19}$$

The series defining f_1 can be extended analytically to all of $\mathbb{D}$, and f_2 can be extended analytically to $\mathbb{C} \setminus \overline{\mathbb{D}}$. We can decompose ϕ now as follows

$$\begin{aligned} \phi(s) &= f(M^{-1}(s)) = f_1(M^{-1}(s)) + f_2(M^{-1}(s)), \\ &=: \phi_1(s) + \phi_2(s), \quad s \in \mathrm{i}\mathbb{R}. \end{aligned} \tag{5.20}$$

Since the Möbius transform and its inverse are analytic, the claims of the analyticity of ϕ_1 and ϕ_2 follow. Note that in the decomposition (5.19) the constant term a_0 was included in the definition of f_1. We could have included a_0 in f_2, but that is the only freedom in the decomposition. Thus ϕ_1 and ϕ_2 are unique up to an additive constant. □

Lemma 5.1.8. *Using the notation of Lemma 5.1.5, let $\phi \in L^\infty(\mathrm{i}\mathbb{R}; B(\mathscr{X}_1, \mathscr{X}_2))$ denote a symbol of H. Then the analytic part of ϕ in $\mathbb{C}_0^+$ (see Lemma 5.1.7) is also a symbol for H, which is determined completely by H up to an additive constant operator.*

Remark 5.1.9. Note that the analytic part of a symbol ϕ in $\mathbb{C}_0^+$ need not necessarily belong to $L^\infty(\mathrm{i}\mathbb{R}; B(\mathscr{X}_1, \mathscr{X}_2))$.

Proof of Lemma 5.1.8: Choose a $L^\infty(\mathrm{i}\mathbb{R}; B(\mathscr{X}_1, \mathscr{X}_2))$ symbol for H, denoted by ϕ, and let ϕ_1 denote the analytic part of ϕ in $\mathbb{C}_0^+$ from Lemma 5.1.7. That ϕ_1 is a symbol of H follows from

$$H_\phi = \mathcal{P}_+ M_{\phi_1+\phi_2} R_C = \mathcal{P}_+ M_{\phi_1} R_C + \mathcal{P}_+ M_{\phi_2} R_C = \mathcal{P}_+ M_{\phi_1} R_C,$$

i.e. the analytic part of ϕ in $\mathbb{C}_0^-$, ϕ_2, plays no role in (5.12). From (5.18) a symbol ϕ of the Hankel operator H is given by

$$\phi(s) = \frac{zg(z)}{2}, \quad s \in \mathrm{i}\mathbb{R},\ z = M^{-1}(s) \in \mathbb{T},$$

where $g \in L^\infty(\mathbb{T})$ satisfies (5.16). Therefore a calculation shows that

$$\phi_1(s) = \sum_{n\in\mathbb{N}_0} \frac{g_{n-1}}{2} z^n, \quad s \in \mathbb{C}_0^+,\ z = M^{-1}(s) \in \mathbb{D},$$

which is completely determined by H (as $g_n = h_n$ for all $n \in \mathbb{N}_0$) apart from the constant term $g_{-1} \in B(\mathscr{X}_1, \mathscr{X}_2)$, as required. □

Definition 5.1.10. Using the notation of Lemma 5.1.5, we define a transfer function corresponding to the Hankel operator H as the analytic part of a symbol in $\mathbb{C}_0^+$ of H, plus an arbitrary constant operator. The set of transfer functions for H is given by

$$\left\{ \phi_1 + D \;\middle|\; D \in B(\mathscr{X}_1, \mathscr{X}_2) \right\},$$

where ϕ_1 denotes the analytic part of a symbol ϕ in $\mathbb{C}_0^+$ of H (see the decomposition (5.20)).

Remark 5.1.11. We remark that transfer functions have now been introduced four times in this thesis, notably on p. 5, Definition 3.1.3, p. 74 and in Definition 5.1.10 above. Each definition has arisen from a slightly different point of view, hence its inclusion, but all are equivalent.

Now that we have defined a transfer function in terms of the Hankel operator, we describe some of the relationships already known in the literature between a Hankel operator and its transfer function(s). For the next result we need to collect various function spaces.

Definition 5.1.12. Let $\mathscr{X}$ denote a Banach space.

$BMO(\mathrm{i}\mathbb{R};\mathscr{X})$**:** The space $BMO(\mathrm{i}\mathbb{R};\mathscr{X})$ is the space of functions from $\mathrm{i}\mathbb{R}$ to $\mathscr{X}$ of bounded mean oscillation, see [25, Section VI] or [65, p. 728]. Specifically, for $f \in L^1_{\mathrm{loc}}(\mathrm{i}\mathbb{R};\mathscr{X})$ and $J \subset \mathbb{R}$ a (finite) interval, the mean oscillation of f over J is

$$\frac{1}{|J|}\int_J \|f(\mathrm{i}\omega) - f_J\|\, d\omega,$$

where f_J is the mean of f on J

$$f_J = \frac{1}{|J|}\int_J f(\mathrm{i}\omega)\, d\omega,$$

and $|J|$ is the Lebesgue measure of J. The space BMO contains those f with uniformly bounded mean oscillation, i.e.

$$\sup_J \frac{1}{|J|}\int_J \|f(\mathrm{i}\omega) - f_J\|\, d\omega < \infty.$$

$BMOA(\mathbb{C}_0^+;\mathscr{X})$**:** Is the set of BMO functions that have an analytic extension to $\mathbb{C}_0^+$.

$VMO(\mathrm{i}\mathbb{R};\mathscr{X})$ The space of functions of vanishing mean oscillation (VMO) is the closed subspace of BMO with the additional property that for J an interval of $\mathbb{R}$

$$\lim_{|J|\to 0}\left[\sup_J \frac{1}{|J|}\int_J \|f(\mathrm{i}\omega) - f_J\|\, d\omega\right] = 0.$$

$VMOA(\mathbb{C}_0^+;\mathscr{X})$: Is the set of VMO functions that have an analytic extension to $\mathbb{C}_0^+$.

$A^p(\mathbb{C}_0^+;\mathscr{X})$, $A^{p,r}(\mathbb{C}_0^+;\mathscr{X})$: For $p > 0$ the Bergman space A^p is the space of analytic functions $f : \mathbb{C}_0^+ \to \mathscr{X}$ such that

$$\int_{x\in\mathbb{R}^+}\int_{y\in\mathbb{R}} \|f(x+\mathrm{i}y)\|^p \, dy\, dx < \infty.$$

The weighted Bergman space $A^{p,r}$ for $r > -\frac{1}{2}$ comprises the analytic functions $f : \mathbb{C}_0^+ \to \mathscr{X}$ such that

$$\int_{x\in\mathbb{R}^+}\int_{y\in\mathbb{R}} \|f(x+\mathrm{i}y)\|^p K(x+\mathrm{i}y, x+\mathrm{i}y)^{-r} \, dy\, dx < \infty,$$

where K is the Bergman kernel (of the right-half plane) given by

$$K(z,w) = \frac{1}{(z+\overline{w})^2}.$$

$B_p(\mathbb{C}_0^+;\mathscr{X})$: The analytic Besov space B_p (with $p > 0$) consists of those analytic functions $f : \mathbb{C}_0^+ \to \mathscr{X}$ which satisfy

$$f^{(n)} \in A^{p,\frac{np}{2}-1}(\mathbb{C}_0^+;\mathscr{X}),$$

for some (and hence any) $n \in \mathbb{N}$ (where $f^{(n)}$ denotes the n^{th} derivative).

$RH^\infty(\mathbb{C}_0^+;\mathscr{X})$: The space RH^∞ consists of rational H^∞ functions.

Theorem 5.1.13. *For Hilbert spaces $\mathscr{X}_1$ and $\mathscr{X}_2$ let $H : L^2(\mathbb{R}^+;\mathscr{X}_1) \to L^2(\mathbb{R}^+;\mathscr{X}_2)$ and G denote a Hankel operator and a member of the set of transfer functions for H respectively. Then*

(i) H is bounded if and only if $G \in BMOA$.

(ii) H is compact if and only if $G \in VMOA$.

Assume additionally that $\mathscr{X}_1$ and $\mathscr{X}_2$ are finite dimensional. Then

(iii) $H \in S_p$ (the Schatten class, see Definition 4.3.2) if and only if $G \in B_p$.

(iv) H is finite rank if and only if $G \in RH^\infty$.

Proof. Note that the additive constant operator in the definition of a transfer function does not matter, so long as it is bounded, see Definition 5.1.10. Part (i) follows from [25, Corollary 6.1.3], which proves that $BMO(\mathbb{T})$ is transformed into $BMO(\mathrm{i}\mathbb{R})$ under the Möbius transform M (and vice versa under M^{-1}), and [65, Theorem 1.1.2].

Similarly, part (*ii*) follows from [65, Theorem 1.5.8] and the same (standard) techniques mapping $BMO(\mathbb{T})$ into $BMO(\mathrm{i}\mathbb{R})$.

Part (*iii*) is based on Peller [65, Theorem 6.1.1] for the scalar case, $\mathscr{Z} = \mathbb{C}$. The general Hilbert space case is considered in [65, Corollary 6.9.4]. The author treats Hankel operators on the unit disc, but as with (*i*) and (*ii*), standard techniques allow the result to be converted to the half-plane case. The author proves that $H \in S_p \iff G \in B_p(\mathbb{C}_0^+; S_p(\mathscr{Z}_1, \mathscr{Z}_2))$. However, $B(\mathscr{Z}_1, \mathscr{Z}_2) = S_p(\mathscr{Z}_1, \mathscr{Z}_2)$ when either $\mathscr{Z}_1$ or $\mathscr{Z}_2$ is finite dimensional.

Part (*iv*) is Kronecker's Theorem for the half-plane, see [63, Corollary 4.9] or [65, Theorem 1.3.2]. □

By Theorem 5.1.13 (*iii*) transfer functions of a nuclear Hankel operator necessarily belong to the Besov space and hence have atomic decompositions given by Coifman & Rochberg [15]. The following result which uses these decompositions is taken from Partington [63, Corollary 7.9].

Corollary 5.1.14. *A Hankel operator* $H : L^2(\mathbb{R}^+; \mathscr{Z}_1) \to L^2(\mathbb{R}^+; \mathscr{Z}_2)$*, where both* $\mathscr{Z}_1$ *and* $\mathscr{Z}_2$ *are finite dimensional, is nuclear if and only if the corresponding transfer function is of the form*

$$G(s) = \sum_{n\in\mathbb{N}} \lambda_n \frac{\operatorname{Re} a_n}{s - a_n}, \tag{5.21}$$

for sequences $(\lambda_n)_{n\in\mathbb{N}} \in \ell^1(B(\mathscr{Z}_1, \mathscr{Z}_2))$ *and* $(a_n)_{n\in\mathbb{N}} \subset \mathbb{C}_0^-$*. The following bound on the* ℓ^1 *norm of* $(\lambda_n)_{n\in\mathbb{N}}$ *holds*

$$\sum_{n\in\mathbb{N}} \|\lambda_n\|_{B(\mathscr{Z}_1,\mathscr{Z}_2)} \le C\|H\|_N, \quad C \text{ a constant}, \tag{5.22}$$

recalling that $\|H\|_N$ *denotes the nuclear norm of* H*, see Definition 4.3.1. The series* (5.21) *converges uniformly on* $Re\, s > 0$*. Furthermore, let* G^p *denote the* p^{th} *partial sum of* (5.21) *with corresponding Hankel operator* H^p*. Then* H^p *converges to* H *in nuclear norm as* $p \to \infty$*.*

Remark 5.1.15. Remember that the transfer function G of a Hankel operator is only determined by the Hankel operator up to the addition of a constant operator. For the transfer function in (5.21) that constant is fixed by the condition

$$\lim_{\substack{s\in\mathbb{R}\\ s\to\infty}} G(s) = 0.$$

We see that the transfer function is regular with zero feedthrough.

Remark 5.1.16. Our main result for Lyapunov balanced truncation Theorem 5.0.2 assumes that the input and output spaces are finite-dimensional. We make this assumption so that we can apply Theorem 5.1.13 (*iii*) and Corollary 5.1.14 above. The decompositions of [15] are only proven in the case when $\mathscr{X}_1$ and $\mathscr{X}_2$ (as in Corollary 5.1.14) are finite-dimensional.

Lemma 5.1.17. *Using the assumptions and notation of Corollary 5.1.14 the function*

$$h(t) := \sum_{n\in\mathbb{N}} \lambda_n(\operatorname{Re} a_n)\mathrm{e}^{a_n t}, \quad t > 0, \tag{5.23}$$

satisfies $h \in L^1(\mathbb{R}^+; B(\mathscr{X}_1, \mathscr{X}_2))$, $\mathcal{L}h = G$, $G \in H^\infty(\mathbb{C}_0^+, B(\mathscr{X}_1, \mathscr{X}_2))$ *and* G *is regular.*

Proof. It is clear from (5.22) that (5.23) converges absolutely and uniformly on $t > 0$. Moreover, h is continuous for $t > 0$. That $h \in L^1$ follows from the Monotone Convergence Theorem applied to the partial sums

$$t \mapsto \sum_{n=1}^{M} \|\lambda_n \operatorname{Re} a_n \mathrm{e}^{a_n t}\| \in L^1.$$

To evaluate $\mathcal{L}h$ let h_m denote the m^{th} partial sum of h. Formally we have for $s \in \mathbb{C}_0^+$

$$\begin{aligned}
(\mathcal{L}h)(s) &= \int_{\mathbb{R}^+} \mathrm{e}^{-st} h(t)\, dt \\
&= \int_{\mathbb{R}^+} \lim_{m\to\infty} \mathrm{e}^{-st} h_m(t)\, dt = \lim_{m\to\infty} \int_{\mathbb{R}^+} \mathrm{e}^{-st} h_m(t)\, dt \qquad (5.24)\\
&= \lim_{m\to\infty} \sum_{n=1}^{m} \lambda_n \frac{\operatorname{Re} a_n}{s - a_n}, \quad \text{by taking the Laplace transform,} \\
&= G(s), \quad \text{from (5.21).}
\end{aligned}$$

It remains to verify the interchanging of the limit and the integral in (5.24) and for this we use the Dominated Convergence Theorem. Observe that for every $m \in \mathbb{N}$ and s with $\operatorname{Re} s > 0$

$$\|\mathrm{e}^{st} h_m(t)\| \le \sum_{n\in\mathbb{N}} \|\lambda_n \operatorname{Re} a_n \mathrm{e}^{a_n t}\|.$$

From the proof of $h \in L^1$, it follows that

$$t \mapsto \sum_{n\in\mathbb{N}} \|\lambda_n \operatorname{Re} a_n \mathrm{e}^{a_n t}\| \in L^1,$$

and hence by Dominated Convergence we infer equality (5.24). Finally, the inequalities

$$\|G(s)\| \le \sup_{t\ge 0} |\mathrm{e}^{-st}| \cdot \|h\|_1 \le \|h\|_1,$$

shows us that $G \in H^\infty$. That G is regular can be proven directly from the series (5.21) or follows from [81, Theorem 5.6.7]. □

Corollary 5.1.18. *Assume that the Hankel operator $H : L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$ is nuclear and both $\mathscr{U}$ and $\mathscr{Y}$ are finite dimensional. Then H is an integral operator given by*

$$(Hf)(t) = \int_0^\infty h(t+s) f(s)\, ds, \quad \forall\, f \in L^2(\mathbb{R}^+; \mathscr{U}),\ a.a\ t \geq 0, \tag{5.25}$$

where h is defined in (5.23) *and in particular belongs to $L^1(\mathbb{R}^+; B(\mathscr{U}, \mathscr{Y}))$.*

Proof. Let H_h denote the integral operator in (5.25), where h is given by (5.23). We show that $H = H_h$. From Lemma 5.1.17 the function h satisfies $\mathcal{L}h = G$, with G given by (5.21) and G a transfer function for H. From [63, Lemma 4.3, Corollary 4.4] it follows that

$$\mathcal{L} H_h \mathcal{L}^{-1} = \mathcal{P}_+ M_G R_C = H_G.$$

We have already seen from Lemma 5.1.5 that

$$\mathcal{L} H \mathcal{L}^{-1} = H_G,$$

and so we deduce the result. □

Remark 5.1.19. In [27, Theorem 2.1] the following bound is proven

$$\|h\|_{L^1} \leq 2\|H\|_N, \tag{5.26}$$

for a nuclear Hankel operator H given by (5.2). The inequality (5.26) would appear to imply the result of Lemma 5.1.17 is redundant. However, the proof of (5.26) in [27] uses the fact that $h \in L^1$, which is an assumption we wanted to avoid (and rather show that it is a consequence of H being nuclear).

Remark 5.1.20. The converse of Corollary 5.1.18 is not true in the sense that the impulse response belonging to L^1 does not imply that the Hankel operator is nuclear. The integral kernel h given by

$$[0, \infty) \ni t \mapsto h(t) := \mathrm{e}^{-t} \chi_{[1,+\infty)}(t),$$

where χ_J is the indicator (also called the characteristic) function on $J \subseteq \mathbb{R}^+$, clearly belongs to L^1. However, the example of Glover, Lam & Partington [28, Example 2.3] shows that the integral Hankel operator with kernel h is not nuclear.

5.2 Convergence of the Schmidt pairs of integral Hankel operators

In this section we describe some of the properties of the Schmidt pairs of a Hankel operator given by (5.2) with L^1 kernel, which by Corollary 5.1.18 includes nuclear Hankel operators. We describe as well some convergence results of the Schmidt pairs when the kernel is approximated in L^1. The key assumption of this section is:

A H is a Hankel[2] operator $L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y})$ given by (5.2) with kernel $h \in L^1(\mathbb{R}^+;B(\mathscr{U},\mathscr{Y}))$. The input space $\mathscr{U}$, and output space $\mathscr{Y}$ are finite dimensional Hilbert spaces.

The next lemma is crucial and we will make frequent use of it.

Lemma 5.2.1. *Let H denote a Hankel operator satisfying* **A**. *Then for $1 \leq p \leq \infty$, H is a bounded operator*

$$L^p(\mathbb{R}^+;\mathscr{U}) \to L^p(\mathbb{R}^+;\mathscr{Y}).$$

We abuse notation and use the symbol H to represent any of these maps. If $\|H\|_p$ denotes the Hankel (operator) norm of a Hankel operator H satisfying **A** *considered as an operator on L^p then*

$$\|H\|_p \leq \|h\|_1. \tag{5.27}$$

Furthermore, if p is finite then H is compact $L^p(\mathbb{R}^+;\mathscr{U}) \to L^p(\mathbb{R}^+;\mathscr{Y})$.

Proof. This is proven in [27, Appendix 1, p. 895]. □

From Lemma 5.2.1 and Remark 4.3.3 it follows that for a Hankel operator H satisfying assumption **A** the singular values of H (viewed as an operator $L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y})$) are precisely the square roots of the countably many eigenvalues of H^*H, which we always order in decreasing magnitude, counted with multiplicities (as in Remark 4.3.4).

Theorem 5.2.2. *Let H denote an operator satisfying assumption* **A** *and let $(\sigma_k)_{k\in\mathbb{N}}$ denote the singular values of H ordered such that $\sigma_k > \sigma_{k+1} \geq 0$ and each with multiplicity $p_k \in \mathbb{N}$. Let $(h_m)_{m\in\mathbb{N}}$ denote a sequence of kernels approximating h in the sense that*

$$h_m \xrightarrow{L^1} h, \quad \text{as} \quad m \to \infty.$$

Define $(H_m)_{m\in\mathbb{N}}$ as the sequence of Hankel operators $L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y})$ given by the integral operators

$$(H_m f)(t) := \int_0^\infty h_m(t+s)f(s)\,ds, \quad \forall\, f \in L^2(\mathbb{R}^+;\mathscr{U}), \quad a.a.\ t \geq 0. \tag{5.28}$$

[2] more precisely (τ_1,τ_2)-Hankel with $\mathscr{U} = \mathscr{X}_1$ and $\mathscr{Y} = \mathscr{X}_2$ in the terminology of Section 5.1.

Let $(\sigma_i^{(m)})_{i\in\mathbb{N}}$ denote the singular values of H_m, also ordered in decreasing magnitude, each with multiplicity $p_i^{(m)}$. Then for all $k \in \mathbb{N}$ there exists $l_k \in \mathbb{N}$ such that with $l_0 := 0$

$$\left.\begin{array}{ll} \sigma_i^{(m)} \to \sigma_k & \text{for } i \in \{l_{k-1}+1, \dots, l_k\} \\ \text{and} & \displaystyle\sum_{i=l_{k-1}+1}^{l_k} p_i^{(m)} \to p_k, \end{array}\right\} \quad \text{as} \quad m \to \infty. \tag{5.29}$$

Choose orthonormal Schmidt pairs of H_m denoted by $(v_{i,r}^{(m)}, w_{i,r}^{(m)})$, where $r \in \{1, 2, \dots, p_i^{(m)}\}$. Then the Schmidt pairs satisfy

$$v_{i,r}^{(m)} \in L^2 \cap W^{1,1}(\mathbb{R}^+; \mathscr{U}), \quad w_{i,r}^{(m)} \in L^2 \cap W^{1,1}(\mathbb{R}^+; \mathscr{Y}), \quad \forall\, i, m \in \mathbb{N},\ \forall\, r.$$

Moreover, there exists a subsequence $(m_j)_{j\in\mathbb{N}}$ along which for each $k \in \mathbb{N}$, all of the p_k Schmidt pairs $(v_{i,r}^{(m_j)}, w_{i,r}^{(m_j)})$, $i \in \{l_{k-1}+1, \dots, l_k\}$, have a L^2 and $W^{1,1}$ limit denoted $(v_{k,q}, w_{k,q})$, for $q \in \{1, 2, \dots, p_k\}$ (note $q = q(r)$ depends on r). Specifically,

$$\left.\begin{array}{ll} v_{i,r}^{(m_j)} \to v_{k,q}, & \text{in } L^2 \text{ and } W^{1,1} \\ w_{i,r}^{(m_j)} \to w_{k,q}, & \text{in } L^2 \text{ and } W^{1,1} \end{array}\right\} \quad \text{as} \quad j \to \infty. \tag{5.30}$$

Moreover, the $(v_{k,q}, w_{k,q})$ are Schmidt pairs for H corresponding to σ_k and $(w_{k,q})_{k\in\mathbb{N}}^{1\le q\le p_k}$ form an orthonormal basis of eigenvectors in L^2 for HH^* and $(v_{k,q})_{k\in\mathbb{N}}^{1\le q\le p_k}$ for H^*H.

Remark 5.2.3. The statement of Theorem 5.2.2 is notation heavy in order to account for the multiplicities of both the singular values $(\sigma_k)_{k\in\mathbb{N}}$ of H and $(\sigma_k^{(m)})_{k\in\mathbb{N}}$ of H_m. The easiest case to understand is when for every $k, m \in \mathbb{N}$ the singular values σ_k and $\sigma_k^{(m)}$ are simple, which is the case restricted to by Glover *et al.* in [27]. The non-simple case is conceptually similar, and we treat it for full generality, although the proofs become more complicated. Intuitively, what is important is that there is a subsequence $(m_j)_{j\in\mathbb{N}}$ along which every sequence $(v_{k,r}^{(m_j)})_{j\in\mathbb{N}}$ and $(w_{k,r}^{(m_j)})_{j\in\mathbb{N}}$ has an L^2 limit (which is also a limit in $W^{1,1}$), and we get "enough" limits, in the sense that the limits of $(v_{k,r}^{(m_j)})_{j\in\mathbb{N}}$ and $(w_{k,r}^{(m_j)})_{j\in\mathbb{N}}$ form an orthonormal basis of eigenvectors for H^*H and HH^* respectively.

Remark 5.2.4. Under the assumptions of Theorem 5.2.2, if additionally all the singular values of H are simple, then the convergence of Schmidt pairs in Theorem 5.2.2 does not require a subsequence. We do not give all the details, but this follows from [27, Appendix 2, p.896] and Lemma 5.2.7 below combined with [12, Exercise 5.5].

Remark 5.2.5. Note that for this section we do not need to assume that our original Hankel operator H is nuclear, instead only that it is an integral operator of the form (5.2) with $h \in L^1$. Certainly, by Corollary 5.1.18, nuclearity of H would imply this.

5.2.1 Proof of Theorem 5.2.2

Here we put together the proof of Theorem 5.2.2. We firstly collect some technical results from Lax [46] and Chatelin [12].

Lemma 5.2.6. *Let $\mathscr{Z}$ denote a Banach space on which is defined a continuous sesquilinear linear form $\langle \cdot, \cdot \rangle$, which induces a new norm under which the completion of $\mathscr{Z}$ is a Hilbert space $\mathscr{H}$. Suppose that T is a bounded operator on $\mathscr{Z}$ such that*

$$\langle Tx, y \rangle = \langle x, Ty \rangle, \quad \forall\, x, y \in \mathscr{Z}.$$

Then

(i) T extends by continuity to an operator in $B(\mathscr{H})$.

(ii) The spectrum of T over $\mathscr{H}$ is a subset of the spectrum of T over $\mathscr{Z}$.

(iii) The point spectrum of T over $\mathscr{Z}$ is contained in the point spectrum of T over $\mathscr{H}$ and the eigenspace of T over $\mathscr{B}$ with respect to an eigenvalue λ is the same as the eigenspace of T over $\mathscr{H}$ with respect to the same eigenvalue.

Proof. See Lax [46]. □

We will make use of a convergence result which is based on [12, Theorem 5.10], which we formulate as the following lemma.

Lemma 5.2.7. *Suppose $T, T_m : \mathscr{Z} \to \mathscr{Z}$ are compact operators on a Banach space $\mathscr{Z}$ such that*

$$T_m \to T, \quad \textit{uniformly as} \quad m \to \infty.$$

Let λ denote an eigenvalue of T and let $(\lambda_m)_{m\in\mathbb{N}}$ denote a sequence of eigenvalues of $(T_m)_{m\in\mathbb{N}}$. If

$$\lambda_m \to \lambda, \quad \textit{as} \quad m \to \infty,$$

and $(v_m)_{m\in\mathbb{N}}$ is a uniformly bounded sequence of eigenvectors corresponding to λ_m, then there exists a subsequence $(v_{m_j})_{j\in\mathbb{N}}$ that converges to an eigenvector of T corresponding to λ.

Proof. The result is a consequence of [12, Theorem 5.10] and [12, Table 5.1]. □

Lemma 5.2.8. *Let $\mathscr{Z}$ denote a Banach space continuously embedded into a Hilbert space $\mathscr{H}$, $\mathscr{Z} \hookrightarrow \mathscr{H}$, so that*

$$\|v\|_{\mathscr{H}} \leq C\|v\|_{\mathscr{Z}}, \quad \forall\, v \in \mathscr{Z}, \tag{5.31}$$

for some constant $C > 0$. Let T, T_m denote compact operators on $\mathscr{X}$ such that

$$T_m \to T, \quad \textit{uniformly on } \mathscr{X} \textit{ as} \quad m \to \infty. \tag{5.32}$$

Fix an eigenvalue λ of T with corresponding eigenvector v. Suppose that there exists a sequence $(\lambda_m)_{m\in\mathbb{N}}$, where λ_m is an eigenvalue of T_m, such that

$$\lambda_m \to \lambda, \quad \textit{as} \quad m \to \infty. \tag{5.33}$$

And suppose also there exist eigenvectors $v^{(m)}$, which are orthonormal in $\mathscr{H}$, of T_m corresponding to λ_m converging to v in the sense that

$$v^{(m)} \xrightarrow{\mathscr{H}} v, \quad \textit{as} \quad m \to \infty. \tag{5.34}$$

Then it follows that along a subsequence

$$v^{(m_j)} \xrightarrow{\mathscr{X}} v, \quad \textit{as} \quad j \to \infty. \tag{5.35}$$

Proof. Fix λ and v some eigenvalue and eigenvector of T. By assumption there exists a sequence $(v_m)_{m\in\mathbb{N}}$ of (orthonormal in $\mathscr{H}$) eigenvectors of T_m such that (5.34) holds. We seek to prove that the convergence in (5.35) holds as well. To that end we define

$$z^{(m)} := \frac{\|v\|_{\mathscr{X}} v^{(m)}}{\|v^{(m)}\|_{\mathscr{X}}}, \quad m \in \mathbb{N}, \tag{5.36}$$

which have

$$\|z^{(m)}\|_{\mathscr{X}} = \|v\|_{\mathscr{X}} < \infty, \quad \forall\, m \in \mathbb{N}. \tag{5.37}$$

Therefore, as all the hypotheses of Lemma 5.2.7 are satisfied, there exists a subsequence (not relabelled) along which

$$z^{(m)} \xrightarrow{\mathscr{X}} \psi, \quad \text{as } m \to \infty, \tag{5.38}$$

with ψ an eigenvector of T. Now the norm equation (5.37) implies that

$$\|\psi\|_{\mathscr{X}} = \|v\|_{\mathscr{X}}, \tag{5.39}$$

and from the continuous embedding (5.31), convergence in (5.38) gives

$$z^{(m)} \xrightarrow{\mathscr{H}} \psi, \quad \text{as } m \to \infty. \tag{5.40}$$

We want to compare the convergence in (5.34) and (5.40), using the definition of $z^{(m)}$

in (5.36). Suppose that $\|v^{(m)}\|_{\mathscr{Z}}$ is unbounded. Then by (5.34) and (5.36) we see that

$$z^{(m)} \xrightarrow{\mathscr{H}} 0, \quad \text{as } m \to \infty, \tag{5.41}$$

and so $\psi = 0$ in $\mathscr{H}$ by uniqueness of limits. Using the injectivity of the inclusion map $\mathscr{Z} \hookrightarrow \mathscr{H}$ we obtain the contradiction

$$0_{\mathscr{Z}} \neq \psi = 0_{\mathscr{Z}}.$$

Therefore $\|v^{(m)}\|_{\mathscr{Z}}$ is bounded and so has a convergent subsequence (not relabelled) with limit $B \geq 0$ given by

$$B := \lim_{m\to\infty} \|v^{(m)}\|_{\mathscr{Z}}.$$

From the continuous embedding (5.31)

$$\|v^{(m)}\|_{\mathscr{Z}} \geq \frac{\|v^{(m)}\|_{\mathscr{H}}}{C} = \frac{1}{C} > 0, \quad \forall\, m \in \mathbb{N},$$

it follows that $B \geq \frac{1}{C} > 0$.

So by definition of $z^{(m)}$ in (5.36) combined with the algebra of limits and the convergence (5.34) we obtain

$$z^{(m)} \xrightarrow{\mathscr{H}} \frac{\|v\|_{\mathscr{Z}} v}{B}, \quad \text{as} \quad m \to \infty,$$

which when compared with (5.40) yields

$$\psi = \frac{\|v\|_{\mathscr{Z}} v}{B}, \quad \text{in } \mathscr{H}. \tag{5.42}$$

Again by the injectivity of the inclusion $\mathscr{Z} \hookrightarrow \mathscr{H}$, from equality (5.42) we infer

$$\psi = \frac{\|v\|_{\mathscr{Z}} v}{B}, \quad \text{in } \mathscr{Z}. \tag{5.43}$$

Taking $\mathscr{Z}$ norms in (5.43) and using (5.39) gives

$$\lim_{m\to\infty} \|v^{(m)}\|_{\mathscr{Z}} =: B = \|v\|_{\mathscr{Z}} \tag{5.44}$$

which when substituted back into (5.43) gives $\psi = v$. Finally from (5.36), (5.38), (5.44) and the algebra of limits we obtain

$$v^{(m)} = \frac{\|v^{(m)}\|_{\mathscr{Z}} z^{(m)}}{\|v\|_{\mathscr{Z}}} \xrightarrow{\mathscr{Z}} \psi = v,$$

as required $\square$

We now consider in more detail some of the properties of a Hankel operator satisfying **A**, seeking to apply the abstract convergence in established in Lemma 5.2.8 to the Schmidt pairs of H and H_m.

Lemma 5.2.9. *Let H denote a Hankel operator satisfying* **A**. *Then H is a compact operator $W^{1,1}(\mathbb{R}^+;\mathscr{U}) \to W^{1,1}(\mathbb{R}^+;\mathscr{Y})$. The Hilbert space adjoint operator H^* satisfies* **A** *with $h(t)$ replaced by $h^*(t)$ and $\mathscr{U}$ and $\mathscr{Y}$ interchanged. Thus Lemma 5.2.1 applies to H^* and moreover H^* is a compact operator $W^{1,1}(\mathbb{R}^+;\mathscr{Y}) \to W^{1,1}(\mathbb{R}^+;\mathscr{U})$. On $L^p(\mathbb{R}^+;\mathscr{Y})$ for $1 \leq p \leq \infty$ and on $W^{1,1}(\mathbb{R}^+;\mathscr{Y})$ the operator H^* is given by*

$$f \mapsto \left(\mathbb{R}^+ \ni t \mapsto (H^*f)(t) = \int_{\mathbb{R}^+} h^*(t+s)f(s)\,ds\right). \tag{5.45}$$

Proof. It follows from Lemma 5.2.1 that H satisfying **A** is a compact operator

$$L^1(\mathbb{R}^+;\mathscr{U}) \to L^1(\mathbb{R}^+;\mathscr{Y}), \quad L^2(\mathbb{R}^+;\mathscr{U}) \to L^2(\mathbb{R}^+;\mathscr{Y}).$$

We now prove that H is a bounded operator $W^{1,1}(\mathbb{R}^+;\mathscr{U}) \to W^{1,1}(\mathbb{R}^+;\mathscr{Y})$ and for this we need the following formula for the derivative of the Hankel operator given by (5.2)

$$\frac{d}{dt}(Hf) = -hf(0) - H\dot{f}, \quad \forall\, f \in W^{1,1}(\mathbb{R}^+;\mathscr{U}). \tag{5.46}$$

Observe that $W^{1,1}$ continuously embeds into the Banach space of continuous functions on $[0,\infty)$ with the supremum norm, and so we can always understand point evaluations (such as $f(0)$ in (5.46)) by choosing a continuous representative. The derivation of (5.46) is given in [27, Appendix 1]. To prove $H : W^{1,1} \to W^{1,1}$ is bounded we consider

$$\|Hf\|_{1,1} = \|Hf\|_1 + \left\|\frac{d}{dt}Hf\right\|_1. \tag{5.47}$$

The first term on the right hand side of (5.47) is clearly bounded by

$$\|H\|_1 \cdot \|f\|_1 \leq \|h\|_1 \cdot \|f\|_{1,1}, \tag{5.48}$$

where we have used the bound (5.27). We now estimate the second term on the right hand side of (5.47). From the formula for the derivative (5.46) we see

$$\begin{aligned}
\left\|\frac{d}{dt}Hf\right\|_1 &\leq \|h(\cdot)f(0)\|_1 + \|H\dot{f}\|_1 \\
&\leq \|h\|_1 \cdot \|f(0)\|_{\mathscr{U}} + \|H\|_1 \cdot \|\dot{f}\|_1 \\
&\leq 2\|h\|_1 \cdot \|f\|_{1,1},
\end{aligned} \tag{5.49}$$

where we have used (5.27) and

$$\|f(0)\|_{\mathscr{U}} \leq \|f\|_\infty \leq \|\dot{f}\|_1 \leq \|f\|_{1,1}. \tag{5.50}$$

Inserting (5.48) and (5.49) into (5.47) we obtain

$$\|Hf\|_{1,1} \leq 3\|h\|_1 \cdot \|f\|_{1,1}, \tag{5.51}$$

and so H defined by (5.2) is well-defined, linear and bounded $W^{1,1} \to W^{1,1}$. To prove compactness let $(f_n)_{n\in\mathbb{N}} \subseteq W^{1,1}$ be a bounded sequence, i.e. there exists a constant $M > 0$ such that

$$\|f_n\|_{1,1} \leq M, \quad \forall\, n \in \mathbb{N}.$$

As such $(f_n)_{n\in\mathbb{N}}$ is a bounded sequence in L^1 and as $H : L^1 \to L^1$ is compact, there exists a convergent and so Cauchy (in L^1) subsequence $(Hf_{\tau_1(n)})_{n\in\mathbb{N}} \subseteq L^1$. The sequence $(\dot{f}_{\tau_1(n)})_{n\in\mathbb{N}}$ is bounded in L^1 and so by the same argument, $(H\dot{f}_{\tau_1(n)})_{n\in\mathbb{N}}$ has a convergent and hence Cauchy subsequence denoted $(H\dot{f}_{\tau_2(n)})_{n\in\mathbb{N}}$.

Observe that the trace map

$$T : W^{1,1}(\mathbb{R}^+;\mathscr{U}) \to \mathscr{U}, \quad Tu = u(0),$$

is bounded and finite rank, so compact. Therefore there exists a subsequence of $(Tf_{\tau_2(n)} = f_{\tau_2(n)}(0))_{n\in\mathbb{N}}$, denoted by $(Tf_{\tau_3(n)})_{n\in\mathbb{N}}$ that is convergent in $\mathscr{U}$ and so Cauchy in $\mathscr{U}$. We now compute for $m, n \in \mathbb{N}$

$$\begin{aligned}
\|Hf_{\tau_3(n)} - Hf_{\tau_3(m)}\|_{1,1} &= \|Hf_{\tau_3(n)} - Hf_{\tau_3(m)}\|_1 \\
&\quad + \left\| \frac{d}{dt}(Hf_{\tau_3(n)} - Hf_{\tau_3(m)}) \right\|_1 \\
&\leq \|Hf_{\tau_3(n)} - Hf_{\tau_3(m)}\|_1 + \|H\dot{f}_{\tau_3(n)} - H\dot{f}_{\tau_3(m)}\|_1 \\
&\quad + \|h\|_1 \cdot \|f_{\tau_3(n)}(0) - f_{\tau_3(n)}(0)\|_{\mathscr{U}},
\end{aligned} \tag{5.52}$$

where we have used the formula (5.46) for the derivative of Hf_{τ_3}. By construction the right hand side of (5.52) converges to zero and so the sequence $(Hf_{\tau_3(n)})_{n\in\mathbb{N}}$ is Cauchy in $W^{1,1}$ and so convergent. This concludes the proof that $H : W^{1,1} \to W^{1,1}$ is compact.

We now focus our attention on the (Hilbert space) adjoint map H^*. Firstly note that $h^* \in L^1(\mathbb{R}^+; B(\mathscr{Y}, \mathscr{U}))$ as

$$\|h^*\|_1 = \int_{\mathbb{R}^+} \|h^*(t)\|_{B(\mathscr{Y},\mathscr{U})}\, dt = \int_{\mathbb{R}^+} \|h(t)\|_{B(\mathscr{U},\mathscr{Y})}\, dt = \|h\|_1. \tag{5.53}$$

A short calculation shows that

$$H^* : L^2(\mathbb{R}^+;\mathscr{Y}) \to L^2(\mathbb{R}^+;\mathscr{U}),$$

is indeed given by (5.45) and so H^* certainly satisfies $\mathbf{A}$ with $\mathscr{U}$ and $\mathscr{Y}$ interchanged and h replaced by h^*. The remaining claims for H^* follow for the reasons as they do for H. $\square$

Lemma 5.2.10. *Let H denote a Hankel operator satisfying $\mathbf{A}$ and define the Banach spaces*

$$\mathscr{Z}_{\mathscr{B}} := L^2 \cap W^{1,1}(\mathbb{R}^+;\mathscr{B}), \quad \|\cdot\|_{\mathscr{Z}_{\mathscr{B}}} := \|\cdot\|_2 + \|\cdot\|_{1,1}, \quad \mathscr{B} \in \{\mathscr{U},\mathscr{Y}\}. \tag{5.54}$$

Then the operators

$$H^*H : \mathscr{Z}_{\mathscr{U}} \to \mathscr{Z}_{\mathscr{U}}, \quad HH^* : \mathscr{Z}_{\mathscr{Y}} \to \mathscr{Z}_{\mathscr{Y}},$$

defined by the composition of (5.2) *with* (5.45) *are compact and every Schmidt pair* (v,w) *of H satisfies*

$$v \in L^2 \cap W^{1,1}(\mathbb{R}^+;\mathscr{U}), \quad w \in L^2 \cap W^{1,1}(\mathbb{R}^+;\mathscr{Y}).$$

Since $W^{1,1}(\mathbb{R}^+) \hookrightarrow C(\mathbb{R}^+)$, i.e. $W^{1,1}$ is continuously embedded in the space of continuous functions on $[0,\infty)$, we will always assume (v,w) are continuous representatives.

Proof. We restrict our attention to $\mathscr{Z}_{\mathscr{U}}$ and H^*H. The proofs for $\mathscr{Z}_{\mathscr{Y}}$ and HH^* are similar. It is straightforward to see that $\mathscr{Z}_{\mathscr{U}}$ is a Banach space. By Lemma 5.2.9 (particularly equations (5.27) and (5.51)) the operators H and H^*

$$H : \mathscr{Z}_{\mathscr{U}} \to \mathscr{Z}_{\mathscr{Y}}, \quad H^* : \mathscr{Z}_{\mathscr{Y}} \to \mathscr{Z}_{\mathscr{U}},$$

and are bounded. Thus the composition H^*H is a bounded operator on $\mathscr{Z}_{\mathscr{U}}$. For compactness, we prove that

$$H : \mathscr{Z}_{\mathscr{U}} \to \mathscr{Z}_{\mathscr{Y}},$$

is compact so that H^*H is compact as the composition of a bounded and compact operator.

To that end choose a sequence $(f_n)_{n\in\mathbb{N}}$ bounded in $\mathscr{Z}_{\mathscr{U}}$. Then $(f_n)_{n\in\mathbb{N}}$ is a bounded sequence in L^2 and as H is compact on L^2 there is a subsequence $(f_{\tau_1(n)})_{n\in\mathbb{N}}$ such that $(Hf_{\tau_1(n)})_{n\in\mathbb{N}}$ is Cauchy in L^2. Additionally, $(f_{\tau_1(n)})_{n\in\mathbb{N}}$ is bounded in $W^{1,1}$ and so as H is compact on $W^{1,1}$, there is another subsequence, denoted $(f_{\tau_2(n)})_{n\in\mathbb{N}}$, such that

$(Hf_{\tau_2(n)})_{n\in\mathbb{N}}$ is convergent and so Cauchy in $W^{1,1}$. Thus

$$\begin{aligned}\|Hf_{\tau_2(n)} - Hf_{\tau_2(l)}\|_{\mathscr{Z}_{\mathscr{U}}} &= \|Hf_{\tau_2(n)} - Hf_{\tau_2(l)}\|_2 + \|Hf_{\tau_2(n)} - Hf_{\tau_2(l)}\|_{1,1} \\ &\to 0, \quad \text{as} \quad n, l \to \infty,\end{aligned}$$

so that $(Hf_{\tau_2(n)})_{n\in\mathbb{N}}$ is Cauchy and therefore convergent in the Banach space $\mathscr{Z}_{\mathscr{Y}}$. Hence $H : \mathscr{Z}_{\mathscr{U}} \to \mathscr{Z}_{\mathscr{Y}}$ is compact.

The claims about the Schmidt pairs now follow from Lemma 5.2.6, with $\mathscr{X} = \mathscr{Z}_{\mathscr{U}}$, $\mathscr{H} = L^2$ and $T = H^*H$. Here we use that the closure of $\mathscr{Z}_{\mathscr{U}}$ in L^2 is L^2, i.e. $\mathscr{Z}_{\mathscr{U}}$ is dense in L^2. □

Lemma 5.2.11. *Let H denote a Hankel operator satisfying* **A** *and choose $(h_m)_{m\in\mathbb{N}}$ such that*

$$h_m \xrightarrow{L^1} h, \quad as \quad m \to \infty. \tag{5.55}$$

Define the operators H_m by (5.28) so that H_m satisfy **A** *with h replaced by h_m and the conclusions of Lemmas 5.2.9 and 5.2.10 hold for H_m, H_m^* and the Schmidt pairs of H_m. Letting $\mathscr{Z}_{\mathscr{U}}$ and $\mathscr{Z}_{\mathscr{Y}}$ denote the Banach spaces from Lemma 5.2.10, there exist constants $C_1, C_2 > 0$ such that*

$$\|H^*H - H_m^*H_m\|_{B(\mathscr{Z}_{\mathscr{U}})} \le C_1\|h - h_m\|_1, \tag{5.56}$$

$$\|HH^* - H_mH_m^*\|_{B(\mathscr{Z}_{\mathscr{Y}})} \le C_2\|h - h_m\|_1. \tag{5.57}$$

Thus $H_m^*H_m$ and $H_mH_m^*$ converge uniformly to H^*H and HH^* respectively as m tends to infinity.

Proof. By definition the operator H_m given by (5.28) with $h_m \in L^1$ satisfies **A** with h replaced by h_m. We now prove the estimate (5.56); the proof of (5.57) is similar. Let $v \in \mathscr{Z}_{\mathscr{U}}$ and consider

$$\|(H^*H - H_m^*H_m)v\|_{\mathscr{Z}_{\mathscr{U}}} = \|(H^*H - H_m^*H_m)v\|_2 + \|(H^*H - H_m^*H_m)v\|_{1,1}. \tag{5.58}$$

The first term on the right hand side of (5.58) is bounded by

$$\begin{aligned}&\|H^*\|_2 \cdot \|H - H_m\|_2 \cdot \|v\|_2 + \|H^* - H_m^*\|_2 \cdot \|H_m\|_2 \cdot \|v\|_2 \\ &\le (\|H\|_2 + \|h_m\|_1) \cdot \|H - H_m\|_2 \cdot \|v\|_2,\end{aligned} \tag{5.59}$$

where we have used the bound (5.27) for H_m, as H_m satisfies **A**. Then for $m \in \mathbb{N}$ sufficiently large invoking (5.27), (5.54) and (5.55) in (5.59) gives

$$(\|H\|_2 + \|h_m\|_1) \cdot \|H - H_m\|_2 \cdot \|v\|_2 \le 3\|h\|_1 \cdot \|h - h_m\|_1 \cdot \|v\|_{\mathscr{B}_{\mathscr{U}}}. \tag{5.60}$$

To bound the second term on the right hand side of (5.58) we use (5.51) and its version for H^*, namely

$$\|H\|_{1,1} \le 3\|h\|_1, \quad \|H^*\|_{1,1} \le 3\|h^*\|_1 = 3\|h\|_1.$$

Applying these bounds gives,

$$\begin{aligned}
&\|H^*\|_{1,1} \cdot \|H - H_m\|_{1,1} \cdot \|v\|_{1,1} + \|H^* - H_m^*\|_{1,1} \cdot \|H_m\|_{1,1} \cdot \|v\|_{1,1} \\
&\le 3(\|h\|_1 + \|h_m\|_1) \cdot \|h - h_m\|_1 \cdot \|v\|_{\mathscr{B}_{\mathscr{U}}}, \\
&\le 3(\|h\|_1 + 2\|h\|_1) \cdot \|h - h_m\|_1 \cdot \|v\|_{\mathscr{B}_{\mathscr{U}}}, \quad m \text{ sufficiently large.} \qquad (5.61)
\end{aligned}$$

Combining (5.60) and (5.61) in (5.58) and collecting gives (5.56). □

We now have all the ingredients to prove Theorem 5.2.2.

Proof of Theorem 5.2.2: The operators H, H_m are compact $L^2(\mathbb{R}^+; \mathscr{U}) \to L^2(\mathbb{R}^+; \mathscr{Y})$ by Lemma 5.2.1. Applying the estimate (5.27) it follows that

$$\|H - H_m\|_2 \le \|h - h_m\|_1,$$

and so H_m converges uniformly to H as $m \to \infty$. Therefore the singular values $\sigma_k^{(m)}$ of H_m with corresponding multiplicities $p_k^{(m)}$ converge as in (5.29) from [60, Lemma 10.19]. Furthermore, from [60, Lemma 10.21], we can choose orthonormal Schmidt pairs of H_m that converge in L^2 to orthonormal Schmidt pairs of H as claimed. It remains to see the $W^{1,1}$ convergence.

The results of Lemmas 5.2.10 and 5.2.11 imply all the hypotheses of Lemma 5.2.8 hold with $\mathscr{H} = L^2$ and $\mathscr{Z} = \mathscr{Z}_{\mathscr{U}}$ (or $\mathscr{Z}_{\mathscr{Y}}$). Therefore we obtain convergence of a subsequence of the Schmidt pairs in $\mathscr{Z}_{\mathscr{U}}$ (or $\mathscr{Z}_{\mathscr{Y}}$), which in particular implies convergence in $W^{1,1}$ as required. We use an induction and diagonal sequence argument to obtain the existence of a single subsequence along which every Schmidt vector converges. Specifically, using the above argument we find a subsequence $(\tau_1(m))_{m\in\mathbb{N}}$ along which

$$\left.\begin{aligned} v_{i,q}^{(\tau_1(m))} &\xrightarrow{L^2, W^{1,1}} v_{1,r}, \\ w_{i,q}^{(\tau_1(m))} &\xrightarrow{L^2, W^{1,1}} w_{1,r} \end{aligned}\right\} \quad \forall\, i \in \{1, 2, \ldots, l_1\}, \forall\, q, r.$$

Using the above argument again we obtain a subsequence of $(\tau_1(m))_{m\in\mathbb{N}}$, denoted $(\tau_2(m))_{m\in\mathbb{N}}$, such that

$$\left.\begin{aligned} v_{i,q}^{(\tau_1(m))} &\xrightarrow{L^2, W^{1,1}} v_{2,r}, \\ w_{i,q}^{(\tau_1(m))} &\xrightarrow{L^2, W^{1,1}} w_{2,r} \end{aligned}\right\} \quad \forall\, i \in \{l_1 + 1, 2, \ldots, l_2\}, \forall\, q, r.$$

By repeating this process we obtain a sequence of subsequences indexed by $\tau_n(m)$, and taking the diagonal sequence $(\tau_m(m))_{m\in\mathbb{N}}$ gives the desired result. $\square$

5.2.2 Relation to earlier work

We briefly explore what consequences the extra assumptions of [27] have on a Hankel operator H satisfying assumption **A**, and its Schmidt pairs. The key difference is whether $h \in L^2$ or not. Remember we have already seen in Example 5.0.1 a system where this is not the case. The following lemma describes some implications of the assumption $h \in L^2$.

Lemma 5.2.12. *Let H denote a Hankel operator satisfying* **A**. *The following are equivalent*

(i) $h \in L^2(\mathbb{R}^+; B(\mathscr{U}, \mathscr{Y}))$.

(ii) $H : W^{1,2}(\mathbb{R}^+;\mathscr{U}) \to W^{1,2}(\mathbb{R}^+;\mathscr{Y})$ *is compact.*

If either (i) or (ii) above hold then

(iii) every Schmidt pair (v, w) of H satisfies

$$v \in W^{1,2}(\mathbb{R}^+;\mathscr{U}), \quad w \in W^{1,2}(\mathbb{R}^+;\mathscr{Y}).$$

Finally if additionally the vectors $(v_{i,k}(0))_{i\in\mathbb{N}}^{1\le k\le p_i}$ span $\mathscr{U}$ then (iii) implies (i).

Remark 5.2.13. The assumption $(v_{i,k}(0))_{i\in\mathbb{N}}^{1\le k\le p_i}$ span $\mathscr{U}$ is always the case if $\mathscr{U}$ is one-dimensional, as for every $i \in \mathbb{N}$ there always exists k such that $v_{i,k}(0) \neq 0$. See [17, Lemma 4.3] for a proof of this assertion when the singular values are simple and [1, Theorem 7.2] for the general case.

Proof of Lemma 5.2.12: $(i) \Rightarrow (ii)$: This is similar to the proof from Lemma 5.2.9 that H is compact on $W^{1,1}$, only now taking L^2 norms instead of L^1 norms. Note that the same formula (5.46) holds for the derivative of Hf, for $f \in W^{1,2}$.

$(ii) \Rightarrow (i)$: Rearranging the derivative formula (5.46) gives

$$h(t)f(0) = -\frac{d}{dt}(Hf)(t) - (H\dot{f})(t), \quad \forall\, f \in W^{1,2}(\mathbb{R}^+;\mathscr{U}). \tag{5.62}$$

The right hand side of (5.62) is in L^2, and hence so is the left hand side. Since $\mathscr{U}$ is finite dimensional, it follows that $h \in L^2$.

(i) or $(ii) \Rightarrow (iii)$: This is analogous to Lemma 5.2.10 and follows in the same way from Lemma 5.2.6.

Now we assume that $(v_{i,k}(0))_{i\in\mathbb{N}}^{1\le k\le p_i}$ span $\mathscr{U}$.

(iii) $\Rightarrow$ *(i)*: That $h \in L^2$ follows readily from the derivative formula for $Hv_{i,k}$, namely

$$h(t)v_{i,k}(0) = -\frac{d}{dt}(Hv_{i,k})(t) + (H\dot{v}_{i,k})(t) = -\sigma_i \dot{w}_{i,k}(t) - H\dot{v}_{i,k}(t).$$

The right hand side is L^2 and thus so is the left hand side. Since this holds on a basis for $\mathscr{U}$ we conclude that $h \in L^2$ as required. $\square$

The significance of the Schmidt vectors belonging to $W^{1,2}$ is that by Lemma 4.1.7, $W^{1,2}$ is the domain of the generator of the semigroup of the output-normal realisation ${}^{sr}\Sigma^2$ from Lemma 4.1.6. This property is used in the balanced truncation of [27] and expanded more upon in Section 5.3.

As we might expect, when $h \in L^2$ and is approximated by h_m in *both* L^1 and L^2, we also get convergence of the Schmidt pairs in $W^{1,2}$ which is described in the following corollary.

Corollary 5.2.14. *Let H denote a Hankel operator satisfying* **A** *and suppose additionally that $h \in L^2$. If $(h_m)_{m\in\mathbb{N}}$ are chosen such that*

$$h_m \xrightarrow{L^1, L^2} h, \quad \text{as} \quad m \to \infty, \tag{5.63}$$

and H_m are given by (5.28) *then all the conclusions of Theorem 5.2.2 hold. Moreover, the choice of Schmidt pairs of H_m in Theorem 5.2.2 converge to the Schmidt pairs of H in $W^{1,2}$* as well *as the senses already established in* (5.30).

Proof. The proof is very similar to that of Theorem 5.2.2. We introduce the Banach spaces

$$\mathscr{Z}'_{\mathscr{B}} := W^{1,2} \cap W^{1,1}(\mathbb{R}^+; \mathscr{B}), \quad \|\cdot\|_{\mathscr{Z}'_{\mathscr{B}}} := \|\cdot\|_{1,2} + \|\cdot\|_{1,1}, \quad \mathscr{B} \in \{\mathscr{U}, \mathscr{Y}\}. \tag{5.64}$$

Again we restrict out attention to $\mathscr{Z}'_{\mathscr{U}}$ and H^*H. Arguing as in Lemma 5.2.10, and also using Lemma 5.2.12, H^*H and $H_m^*H_m$ are compact on $\mathscr{Z}'_{\mathscr{U}}$. Furthermore, a calculation shows that there exists constants $C_3, C_4 > 0$ such that

$$\|H^*H - H_m^*H_m\|_{B(\mathscr{Z}'_{\mathscr{U}})} \leq C_3\|h - h_m\|_1 + C_4\|h - h_m\|_2.$$

By our assumption (5.63) it follows that $H_m^*H_m$ converges uniformly to H^*H on $\mathscr{B}'_{\mathscr{U}}$ as m tends to infinity.

The $W^{1,1}$ and $W^{1,2}$ convergence now follows from Lemma 5.2.8 with $\mathscr{H} = L^2$ and $\mathscr{Z} = \mathscr{Z}'_{\mathscr{U}}$ (as convergence in $\mathscr{Z}'_{\mathscr{U}}$ implies convergence in $W^{1,1}$ and $W^{1,2}$ via (5.64)). $\square$

Remark 5.2.15. In [27] the authors choose sequences of partial sums of the Coifman & Rochberg decompositions, see Proposition 5.1.14, as approximations of h and G. This guarantees nuclear convergence of the Hankel operators, and so L^1 convergence of the kernels and H^∞ convergence of the transfer functions (see the inequalities (5.125) for a proof of these assertions). In [27, Lemma 4.2], the authors tweak the approximating sequence G^m by setting

$$F^m(s) := \frac{G^m(s)}{1+\varepsilon_m s}, \quad m \in \mathbb{N},$$

for some sequence of positive numbers $(\varepsilon_m)_{m\in\mathbb{N}}$ converging to zero. The sequence $(F^m)_{m\in\mathbb{N}}$ converges to G in the above senses, but also in H^2. Therefore the impulse responses converge in L^1 and L^2. We remark in Section 5.3.3 how L^2 convergence of the impulse responses is used in [27]. We remark here though that in light of Corollary 5.2.14, it follows that the Schmidt pairs of the Hankel operators corresponding to F_m converge in $W^{1,2}$ to those of the Hankel operator corresponding to G.

Remark 5.2.16. In [27] the space of absolutely continuous, uniformly bounded functions with distributional derivatives in L^1 is used and denoted by C^1. This space C^1 is also used by Adamjan *et al.* in [1]. Here C^1 is equipped with the norm

$$\|f\|_{C^1} := \|f\|_\infty + \|\dot{f}\|_1.$$

A short calculation shows $W^{1,1} \hookrightarrow C^1$, i.e. $W^{1,1}$ is continuously embedded into C^1. As such we recover from Lemma 5.2.10 that the Schmidt pairs belong to C^1. Additionally, under the assumptions of Theorem 5.2.2, from that result we see that the Schmidt pairs of H_m converge in C^1 to Schmidt pairs of H.

We have chosen to use $W^{1,1}$ instead of C^1 because of Lemma 4.1.7, which gives that $W^{1,1}$ is the domain of the generator of the semigroup of the exactly observable shift realisation ${}^{sr}\Sigma^1$ from Lemma 4.1.6. This will become important for defining truncated systems.

5.3 Realisations and truncated realisations of integral Hankel operators

In this section we define the reduced order system obtained by Lyapunov balanced truncation of a system with bounded Hankel satisfying assumption **A** from Section 5.2. For this we first need a realisation of such a Hankel operator. As described in Chapter 4, the term realisation is usually understood as a realisation of an input-output map on L^p (equivalently of a transfer function). However, by [81, Theorem 5.6.7] a system with impulse response $h \in L^1(\mathbb{R}^+; B(\mathscr{U};\mathscr{Y}))$ has a transfer function which is regular in the uniform topology with zero feedthrough. Recall from Section 5.1 that the transfer function is only determined by the Hankel operator up to an additive constant, the feedthrough. By ensuring $h \in L^1$ we have fixed zero feedthrough and so the Hankel operator completely determines the transfer function. Therefore for the class of systems we consider, a realisation of the transfer function is equivalent to a realisation of the Hankel operator.

The aim of this section is to prove Theorem 5.0.3 and also provide the ingredients to prove Theorem 5.0.2. We prove Theorem 5.0.3 by finding a realisation of G_n^m that converges to a realisation of G_n. We have a similar strategy to Glover & Curtain [17] and Glover *et al.* [27], in that we seek realisations that we can describe in terms of the Schmidt pairs of the Hankel operators. Our novel approach is then to use the $W^{1,1}$ convergence of the Schmidt pairs established in Section 5.2. Propositions 5.3.9 and 5.3.11 are the main results of this section; the former is a more detailed version of Theorem 5.0.3 and describes convergence properties of approximate balanced truncations to the exact balanced truncation. The latter describes some properties of the reduced order system.

We remind the reader that a version of Theorem 5.0.3 has been proven for a specific approximation in [27] under the stronger assumptions listed on p. 79. We are only assuming that the Hankel operator satisfies **A** from Section 5.2 and seek to derive convergence in H^∞ of *any* approximate sequence of reduced order transfer functions G_n^m, satisfying $h_m \xrightarrow{L^1} h$, to the exact reduced order transfer function G_n.

We realise a bounded Hankel operator H satisfying assumption **A** via the exactly observable shift realisation ${}^{sr}\Sigma^1$ from Lemma 4.1.6. In the next lemma we describe the generators of this realisation.

Lemma 5.3.1. *Let H denote a Hankel operator satisfying* **A***, with transfer function G. Then for $1 \le p < \infty$ the shift realisation ${}^{sr}\Sigma^p$ of G from Lemma 4.1.6 is an L^p well-posed linear realisation of H and has generators A, B and C given by*

$$A : D(A) \to L^p(\mathbb{R}^+;\mathscr{Y}), \qquad A = \frac{d}{dt}, \quad D(A) = W^{1,p}(\mathbb{R}^+;\mathscr{Y}), \tag{5.65}$$

$$B : \mathscr{U} \to W^{-1,p}(\mathbb{R}^+;\mathscr{Y}), \qquad (Bu)(t) = h(t)u, \quad p > 1, \tag{5.66}$$

$$C : D(A) \to \mathscr{Y} \qquad Cx = x(0). \tag{5.67}$$

where $W^{-1,p}(\mathbb{R}^+;\mathscr{Y})$ is the dual of $W_0^{1,p}(\mathbb{R}^+;\mathscr{Y})$. When $p = 1$ the control operator B is bounded and is defined by

$$B : \mathscr{U} \to L^1(\mathbb{R}^+;\mathscr{Y}), \qquad (Bu)(t) = h(t)u, \tag{5.68}$$

Proof. The main operator A was described in Lemma 4.1.7. By [81, Example 4.4.6] the operator C in (5.67) is the observation operator of ${}^{sr}\Sigma^p$. To find the control operator we consider first the case $p > 1$. Note that since $\mathscr{X} := L^p$ is reflexive for $p > 1$ it follows as in [81, Remark 3.6.1] that $(L^p)_{-1} = \mathscr{X}_{-1} = D(A^*)' = (W_0^{1,p})' = W^{-1,p}$, where the subscript -1 denotes the usual rigged space. We now claim that the map B in (5.66) is well-defined. Observe that $Bu \in L^1(\mathbb{R}^+;\mathscr{Y})$. We claim that every element of L^1 gives rise to an element of $W^{-1,p}$. For $f \in L^1$ define

$$W_0^{1,p} \ni \phi \mapsto T_f[\phi] = \int_0^\infty \langle f(s), \phi(s) \rangle_{\mathscr{Y}} \, ds,$$

which is certainly linear and complex valued. It remains to see that T_f is bounded. Let $C^{k,\gamma}(\mathbb{R})$ denote the Hölder space, with norm $\| \cdot \|_{C^{k,\gamma}}$, see, for example, Evans [24, p.241]. For $\phi \in W_0^{1,p}(\mathbb{R}^+; \mathscr{Y})$ we have for $\alpha := 1 - \frac{1}{p} \in (0, 1)$

$$\|\phi\|_\infty \leq \|\phi\|_{C^{0,\alpha}} \leq C\|\phi\|_{1,p}, \tag{5.69}$$

for some constant $C > 0$ (depending only on p), where the second inequality is Morrey's inequality, viewing elements of $W_0^{1,p}(\mathbb{R}^+; \mathscr{Y})$ as belonging to $W^{1,p}(\mathbb{R}, \mathscr{Y})$ by extending by zero. Therefore combining the Hölder inequality with (5.69) gives

$$|T_f[\phi]| \leq \|f\|_1 \cdot \|\phi\|_\infty \leq C\|f\|_1 \cdot \|\phi\|_{1,p},$$

and hence $T_f \in W^{-1,p}$. So now the input map of the system with generators (A, B) is given by

$$\int_{\mathbb{R}^-} \tau^{-s} B u(s) \, ds,$$

which we can either prove is well-defined directly, or by using that B maps into $\mathscr{X}_{-1}$, which was established above. Using the formula (5.66) we see that for $t \geq 0$ the above input map satisfies

$$\int_{\mathbb{R}^-} \tau^{-s} h(t) u(s) \, ds = \int_{\mathbb{R}^+} h(t+s) u(-s) \, ds = (HRu)(t),$$

i.e. the above input map is equal to the reflected Hankel operator HR, which is the input map of ${}^{sr}\Sigma^1$. By the uniqueness of a control operator in $B(\mathscr{U}, (L^p(\mathbb{R}^+; \mathscr{Y}))_{-1})$, B defined by (5.66) must be the control operator for ${}^{sr}\Sigma^1$.

The proof of the case $p = 1$ is simpler, as now B maps into $\mathscr{X}$ instead of $\mathscr{X}_{-1}$, and just repeats the last part of the above proof. □

Remark 5.3.2. The exactly observable shift realisation ${}^{sr}\Sigma^1$ is generally not approximately controllable, and so not minimal. However, by [81, Theorem 9.1.9 (i)] we can obtain a minimal realisation from ${}^{sr}\Sigma^1$ by changing (reducing) the state space to $\overline{\operatorname{im} HR}$, the reachable subspace, instead. That ${}^{sr}\Sigma^1$ is not necessarily controllable is not an issue, as we will see in Section 5.3.1 that the truncation method gives rise to a minimal finite-dimensional system.

We need the following "adjoint" operators to those of Lemma 5.3.1.

Lemma 5.3.3. *Let H denote a Hankel operator satisfying* $\mathbf{A}$ *and let A, B denote the* *generating operators from Lemma 5.3.1 with $p = 1$. Then the operators defined by*

$$
\begin{aligned}
&A^* : D(A^*) \to L^1(\mathbb{R}^+;\mathscr{Y}), && A^* = -\frac{d}{dt}, \quad D(A^*) = W_0^{1,1}(\mathbb{R}^+;\mathscr{Y}),\\
&B^* : D(A) \to \mathscr{U}, && B^*x = (H^*x)(0),
\end{aligned}
\tag{5.70}
$$

are adjoint to A and B in the sense that

$$
\begin{aligned}
\langle Ax, y\rangle_{L^2} &= \langle x, A^*y\rangle_{L^2}, \quad \forall x \in D(A),\ \forall\, y \in D(A^*),\\
\langle Bu, x\rangle_{L^2} &= \langle u, B^*x\rangle_{\mathscr{U}}, \quad \forall u \in \mathscr{U},\ \forall\, x \in D(A).
\end{aligned}
\tag{5.71}
$$

The above L^2 inner products are understood as the duality pairing of L^1 and L^∞ (the latter containing $W^{1,1}$). Recall here that $D(A) = W^{1,1}(\mathbb{R}^+;\mathscr{Y})$.

Proof. For the adjoint property (5.71) between A and A^* the key calculation is

$$
\langle Ax, y\rangle_{L^2} = \langle \dot{x}, y\rangle_{L^2} = [\langle x(t), y(t)\rangle_{\mathscr{Y}}]_0^\infty - \langle x, \dot{y}\rangle_{L^2},
$$

where we have integrated by parts. Now using that $x \in W^{1,1}$

$$
\begin{aligned}
\langle Ax, y\rangle_{L^2} &= -\langle x(0), y(0)\rangle_{\mathscr{Y}} - \langle x, \dot{y}\rangle_{L^2},\\
&= -\langle x, \dot{y}\rangle_{L^2} = \langle x, A^*y\rangle_{L^2},
\end{aligned}
$$

when $y \in D(A^*) = W_0^{1,1}(\mathbb{R}^+;\mathscr{Y})$. We now consider B^*, which is certainly well-defined on its domain as for $x \in D(A)$

$$
\|B^*x\|_{\mathscr{U}} = \|(H^*x)(0)\|_{\mathscr{U}} \le \|H^*x\|_\infty \le \|h^*\|_1 \cdot \|x\|_\infty \le \|h\|_1 \cdot \|x\|_{1,1},
$$

where we have used the bound (5.27) for H^*. To see the adjoint property (5.71) observe that

$$
\begin{aligned}
\langle Bu, x\rangle_{L^2} &= \int_{\mathbb{R}^+} \langle h(s)u, x(s)\rangle_{\mathscr{Y}}\, ds = \left\langle u, \int_{\mathbb{R}^+} h^*(s)x(s)\, ds \right\rangle_{\mathscr{U}}\\
&= \langle u, (H^*x)(0)\rangle_{\mathscr{U}} = \langle u, B^*x\rangle_{\mathscr{U}}.
\end{aligned}
$$

□

5.3.1 Truncations of the exactly observable shift realisation

The Lyapunov balanced truncation is defined in Definition 5.3.5 below. Before that we need the following technical result.

Lemma 5.3.4. *Let $(w_{i,k})_{i\in\mathbb{N}}^{1\le k\le p_i}$ denote an orthonormal basis of Schmidt vectors of a*

Hankel operator satisfying **A**. *Then for* $n \in \mathbb{N}$ *define*

$$\mathscr{X}_n := \langle w_{i,k} \,|\, 1 \le i \le n,\ 1 \le k \le p_i \rangle, \tag{5.72}$$

which is a closed subspace of L^1, $W^{1,1}$ *and* L^2. *We use the notation* $\mathscr{X}_n^1$, $\mathscr{X}_n^{1,1}$, *and* $\mathscr{X}_n^2$ *to denote* $\mathscr{X}_n$ *considered as a subspace of* L^1, $W^{1,1}$ *and* L^2 *respectively. Then there exist complementary subspaces* $\mathscr{Z}_n^1$, $\mathscr{Z}_n^{1,1}$, *and* $\mathscr{Z}_n^2$ *such that*

$$\begin{aligned} L^1(\mathbb{R}^+;\mathscr{Y}) &= \mathscr{X}_n^1 \oplus \mathscr{Z}_n^1, \\ W^{1,1}(\mathbb{R}^+;\mathscr{Y}) &= \mathscr{X}_n^{1,1} \oplus \mathscr{Z}_n^{1,1}, \\ L^2(\mathbb{R}^+;\mathscr{Y}) &= \mathscr{X}_n^2 \oplus \mathscr{Z}_n^2, \end{aligned}$$

and these decompositions are all orthogonal with respect to the L^2 *inner product or duality product as appropriate. There also exist continuous projections*

$$\begin{aligned} &\mathbb{P}_n : L^2(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n^2, && \mathbb{Q}_n := I - \mathbb{P}_n : L^2(\mathbb{R}^+) \to \mathscr{Z}_n^2, \\ &P_n : W^{1,1}(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n^{1,1}, && Q_n := I - P_n : W^{1,1}(\mathbb{R}^+) \to \mathscr{Z}_n^{1,1}, \\ &\mathcal{P}_n : L^1(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n^1, && \mathcal{Q}_n := I - \mathcal{P}_n : L^1(\mathbb{R}^+) \to \mathscr{Z}_n^1. \end{aligned}$$

P_n *is a restriction of* $\mathbb{P}_n$ *and* $\mathcal{P}_n$ *is the continuous extension of* P_n. *Each projection* $\mathbb{P}_n, P_n$ *and* $\mathcal{P}_n$ *is given by*

$$x \mapsto \sum_{i=1}^{n} \sum_{k=1}^{p_i} \langle w_{i,k}, x \rangle_{L^2} w_{i,k} \tag{5.73}$$

on its domain. The projections $\mathcal{P}_n, \mathcal{Q}_n$, P_n *and* Q_n *satisfy*

$$\left.\begin{aligned} \langle x, \mathcal{P}_n y \rangle_{L^2} &= \langle \mathcal{P}_n x, y \rangle_{L^2}, \\ \langle x, \mathcal{Q}_n y \rangle_{L^2} &= \langle \mathcal{Q}_n x, y \rangle_{L^2} \end{aligned}\right\} \quad \forall\, x, y \in L^1(\mathbb{R}^+;\mathscr{Y}), \tag{5.74}$$

$$\left.\begin{aligned} \langle x, P_n y \rangle_{L^2} &= \langle P_n x, y \rangle_{L^2}, \\ \langle x, Q_n y \rangle_{L^2} &= \langle Q_n x, y \rangle_{L^2} \end{aligned}\right\} \quad \forall\, x, y \in W^{1,1}(\mathbb{R}^+;\mathscr{Y}). \tag{5.75}$$

Equation (5.74) *is understood as the duality-pairing.*

Proof. The proof is reasonably long, but elementary, and does not particularly contribute to the understanding of the material presented in this chapter. As such we have placed the proof in Appendix B. □

Definition 5.3.5. Let (A, B, C) denote the generating operators from Lemma 5.3.1 of the L^1 well-posed linear system ${}^{sr}\Sigma^1$ realising a Hankel operator satisfying assumption **A**. Using the decompositions and projections of Lemma 5.3.4 for $n \in \mathbb{N}$ define the operators

$$A_n := \mathcal{P}_n A|_{\mathscr{X}_n^{1,1}}, \quad B_n := \mathcal{P}_n B, \quad C_n := C|_{\mathscr{X}_n^{1,1}}. \tag{5.76}$$

The operators in (5.76) generate a finite-dimensional linear system on $(\mathscr{Y}, \mathscr{X}_n, \mathscr{U})$, called the reduced order system obtained by Lyapunov balanced truncation, or just the Lyapunov balanced truncation, which we denote by $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$. The function

$$G_n(s) := C_n(sI - A_n)^{-1}B_n, \tag{5.77}$$

defined and analytic on some right-half plane, is the called the reduced order transfer function obtained by Lyapunov balanced truncation.

Remark 5.3.6.

1. For the operators defined in (5.76) to make sense it is crucial that $\mathscr{X}_n \subseteq D(A)$ and that B is bounded, which was established in Lemmas 5.2.10 and 5.3.1 respectively.
2. In Lemma 5.3.4 we define $\mathscr{X}_n$ as the direct sum of eigenspaces of H^*H corresponding to the first n eigenvalues, which throughout this work are the n *largest* eigenvalues. Recall the square roots of the eigenvalues of H^*H are the singular values of H. Keeping the largest singular values in the truncated system, and omitting the rest, is essential for a tighter error bound in (5.7). In principle, however, we could define a truncated system as in Definition 5.3.5 by restricting and projecting onto *any* sum of eigenspaces.

Remark 5.3.7. We now drop the distinction $\mathscr{X}_n^1$, $\mathscr{X}_n^{1,1}$, $\mathscr{X}_n^2$ and simply consider A_n as an operator

$$A_n : \mathscr{X}_n \to \mathscr{X}_n,$$

where $\mathscr{X}_n$ is still given by (5.72) and is equipped with the L^2 inner product, so that $(\mathscr{X}_n, \|\cdot\|_2)$ is a finite-dimensional Hilbert space.

That A_n is bounded on $\mathscr{X}_n$ follows from norm equivalence of norms on a finite dimensional space, although it can also be proved directly.

Remark 5.3.8. Let $\mathtt{u} := \dim \mathscr{U}$ and $\mathtt{y} := \dim \mathscr{Y}$ and choose orthonormal bases $(y_i)_{i=1}^{\mathtt{y}}$, $(w_{i,k})_{1\le i\le n}^{1\le k\le p_i}$ and $(u_i)_{i=1}^{\mathtt{u}}$ for $\mathscr{Y}$, $\mathscr{X}_n$ and $\mathscr{U}$ respectively. Then the operators (A_n, B_n, C_n) in (5.76) have (block) matrix representations with respect to the above bases:

$$\begin{aligned}
\mathcal{A}(n) &:= (\mathcal{A}_{ij})_{i,j=1}^n, & \mathcal{A}_{ij} &\in \mathbb{C}^{p_i\times p_j}, & (\mathcal{A}_{ij})_{kl} &= \langle w_{i,k}, \dot{w}_{j,l}\rangle_{L^2}, \\
\mathcal{B}(n) &:= (\mathcal{B}_i)_{i=1}^n, & \mathcal{B}_i &\in \mathbb{C}^{p_i\times \mathtt{u}}, & (\mathcal{B}_i)_{kl} &= \langle \sigma_i v_{i,k}(0), u_l\rangle_{\mathscr{U}}, \\
\mathcal{C}(n) &:= (\mathcal{C}_i)_{i=1}^n, & \mathcal{C}_i &\in \mathbb{C}^{\mathtt{y}\times p_i}, & (\mathcal{C}_i)_{kl} &= \langle y_k, w_{i,l}(0)\rangle_{\mathscr{Y}}.
\end{aligned} \tag{5.78}$$

5.3.2 Properties of the balanced truncation and Lyapunov equations

We now have the ingredients to state and prove the two main results of this section, Proposition 5.3.9 and Proposition 5.3.11. Both of these results are used in the proof of the error bound of Theorem 5.0.2. Proposition 5.3.9 is a more detailed version of Theorem 5.0.3 and describes convergence properties of approximate balanced truncations to the exact balanced truncation.

Proposition 5.3.9. *Let H denote a Hankel operator satisfying assumption* **A** *with transfer function G. Choose orthonormal bases $(y_i)_{i=1}^{\boldsymbol{y}}$ and $(u_i)_{i=1}^{\boldsymbol{u}}$ for $\mathscr{Y}$ and $\mathscr{U}$ respectively, where $\boldsymbol{y} = \dim \mathscr{Y}$ and $\boldsymbol{u} = \dim \mathscr{U}$. Let $(h_m)_{m\in\mathbb{N}}$ denote any sequence of kernels in $L^1(\mathbb{R}^+; B(\mathscr{U};\mathscr{Y}))$, chosen such that*

$$h_m \xrightarrow{L^1} h, \quad m \to \infty.$$

Define the sequence of transfer functions $(G^m := \mathcal{L}h_m)_{m\in\mathbb{N}}$. Let (A^m, B^m, C^m) denote the generators from Lemma 5.3.1 of the exactly observable shift realisation on L^1 of G^m. For $n \in \mathbb{N}$ let (A_n^m, B_n^m, C_n^m) denote the Lyapunov balanced truncation of G^m from Definition 5.3.5 on $(\mathscr{Y}, \mathscr{X}_n^m, \mathscr{U})$. If the Schmidt vectors defining $\mathscr{X}_n^m$ are chosen as in Theorem 5.2.2 then there exists a subsequence $(\tau(s))_{s\in\mathbb{N}}$, such that

(i) the matrix representations of $A_n^{\tau(s)}, B_n^{\tau(s)}$ and $C_n^{\tau(s)}$ with respect to the bases

$$(y_i)_{i=1}^{\boldsymbol{y}}, \quad (w_{i,k}^{(\tau(s))})_{1\le i\le l_n}^{1\le k\le p_i^{(\tau(s))}} \quad \text{and} \quad (u_i)_{i=1}^{\boldsymbol{u}} \quad \text{for } \mathscr{Y},\ \mathscr{X}_n^{\tau(s)} \text{ and } \mathscr{U}$$

converge element wise to matrix representations of A_n, B_n and C_n with respect to the bases

$$(y_i)_{i=1}^{\boldsymbol{y}}, \quad (w_{i,k})_{1\le i\le n}^{1\le k\le p_i} \quad \text{and} \quad (u_i)_{i=1}^{\boldsymbol{u}} \quad \text{for } \mathscr{Y},\ \mathscr{X}_n \text{ and } \mathscr{U}.$$

The operators A_n, B_n and C_n are truncated operators from Definition 5.3.5 of the exactly observable shift realisation on L^1 of H.

(ii)

$$G_n^{\tau(s)} \xrightarrow{H^\infty} G_n, \quad \text{as} \quad s \to \infty, \tag{5.79}$$

where $G_n^{\tau(s)}$ and G_n are the reduced order transfer functions obtained by Lyapunov balanced truncation from $G^{\tau(s)}$ and G respectively.

Remark 5.3.10. Under the assumptions Proposition 5.3.9, if additionally the singular values of H are simple then all the convergence in Proposition 5.3.9 holds without needing a subsequence.

Proof of Proposition 5.3.9: From Theorem 5.2.2 there exists a subsequence along which every Schmidt vector of H^m converges in L^2 and $W^{1,1}$ to a Schmidt vector of H. For notational convenience within this proof, we denote the terms of the subsequence by m. Using the notation of Theorem 5.2.2, we describe the convergence as $\forall\, k \in \mathbb{N}$,

$\forall\, i \in \{l_{k-1}+1, \dots, l_k\}$

$$\begin{aligned} v_{i,r}^{(m)} &\xrightarrow{W^{1,1}} v_{k,q}, \\ w_{i,r}^{(m)} &\xrightarrow{W^{1,1}} w_{k,q}, \end{aligned} \quad \text{as } m \to \infty,\ q \in \{1, 2, \dots, p_k\} \tag{5.80}$$

Furthermore, from Theorem 5.2.2 for $n \in \mathbb{N}$ and $1 \le k \le n$, $1 \le i \le p_k$ the $w_{i,k}$ form an orthonormal (in L^2) basis for $\mathscr{X}_n$ given by (5.72). Define matrices $(\mathcal{A}(n), \mathcal{B}(n), \mathcal{C}(n))$ by (5.78) in Remark 5.3.8, with entries in terms of the above $W^{1,1}$ limits. Since these $W^{1,1}$ limits are Schmidt pairs of H, it follows that $(\mathcal{A}(n), \mathcal{B}(n), \mathcal{C}(n))$ are the matrix representations (with respect to the bases $(y_i)_{i=1}^{\mathtt{y}}$, $(w_{i,k})_{1\le i\le n}^{1\le k\le p_i}$ and $(u_i)_{i=1}^{\mathtt{u}}$) of A_n, B_n and C_n respectively.

We let $(\mathcal{A}^m(n), \mathcal{B}^m(n), \mathcal{C}^m(n))$ denote the matrix representations of the truncation (A_n^m, B_n^m, C_n^m) with respect to the bases $(y_i)_{i=1}^{\mathtt{y}}$, $(w_{i,k}^{(m)})_{1\le i\le l_n}^{1\le k\le p_i^{(m)}}$ and $(u_i)_{i=1}^{\mathtt{u}}$. These matrices are given by

$$\mathcal{A}^m(n) = (\mathcal{A}_{ij}^m)_{i,j=1}^{l_n}, \quad \mathcal{B}^m(n) = (\mathcal{B}_i^m)_{i=1}^{l_n}, \quad \mathcal{C}^m(n) = (\mathcal{C}_i^m)_{i=1}^{l_n}, \tag{5.81}$$

where $\mathcal{A}_{ij}^m$, $\mathcal{B}_i^m$ and $\mathcal{C}_i^m$ are as in (5.78) (with entries in terms of the Schmidt vectors of H_m). We use this notation to accommodate for the multiplicities of the singular values $\sigma_k^{(m)}$. We prove the following convergence for every $i, j \in \{1, 2, \dots, n\}$

$$\mathcal{A}_{ij} = \lim_{m\to\infty} \begin{bmatrix} \mathcal{A}^m_{l_{i-1}+1, l_{j-1}+1} & \cdots & \mathcal{A}^m_{l_{i-1}+1, l_j} \\ \vdots & \ddots & \vdots \\ \mathcal{A}^m_{l_i, l_{i-j}+1} & \cdots & \mathcal{A}^m_{l_i, l_j} \end{bmatrix}, \tag{5.82}$$

$$\mathcal{B}_i = \lim_{m\to\infty} \begin{bmatrix} \mathcal{B}^m_{l_{i-1}+1} \\ \vdots \\ \mathcal{B}^m_{l_i} \end{bmatrix}, \tag{5.83}$$

$$\mathcal{C}_i = \lim_{m\to\infty} \left[\mathcal{C}^m_{l_{i-1}+1}, \dots, \mathcal{C}^m_{l_i} \right], \tag{5.84}$$

where the convergence is considered component wise. Equations (5.82)-(5.84) imply that the convergence in (i) holds. So we seek to verify (5.82)-(5.84) and to that end note that for m sufficiently large we have

$$p_i = \sum_{\kappa=l_{i-1}+1}^{l_i} p_\kappa^{(m)}, \quad p_j = \sum_{\kappa=l_{j-1}+1}^{l_j} p_\kappa^{(m)},$$

by (5.29) in Theorem 5.2.2. Thus the matrices on either side of the equality (5.82) are the same size as each other, and similarly for (5.83)-(5.84). To see the component wise convergence, fix

$$\begin{aligned}
p &\in \{l_{i-1}+1,\ldots,l_i\}, \quad a \in\{1,\ldots,p_i^{(m)}\} \\
q &\in \{l_{j-1}+1,\ldots,l_j\}, \quad b \in\{1,\ldots,p_j^{(m)}\}.
\end{aligned} \tag{5.85}$$

We prove (5.82) first. With p,q,a,b as in (5.85) we have

$$(\mathcal{A}^m_{pq})_{ab} = \left\langle w^{(m)}_{p,a}, \dot{w}^{(m)}_{q,b} \right\rangle_{L^2}, \tag{5.86}$$

is an entry of the matrix on the right hand side of (5.82), say the $(\alpha,\beta)^{th}$. By construction

$$\langle w_{i,r_a}, \dot{w}_{j,r_b} \rangle_{L^2}, \tag{5.87}$$

is the $(\alpha,\beta)^{th}$ entry of $\mathcal{A}_{ij}$. To prove (5.82) we prove that (5.86) converges to (5.87) as $m\to\infty$. We have

$$\begin{aligned}
&\left| \langle w_{i,r_a}, \dot{w}_{j,r_b} \rangle_{L^2} - \left\langle w^{(m)}_{p,a}, \dot{w}^{(m)}_{q,b} \right\rangle_{L^2} \right| \\
&\leq \|w_{i,r_a}\|_\infty \|\dot{w}_{j,r_b} - \dot{w}^{(m)}_{q,b}\|_1 + \|w_{i,r_a} - w^{(m)}_{p,a}\|_\infty \|\dot{w}^{(m)}_{q,b}\|_1,
\end{aligned}$$

by the triangle and Hölder inequalities. Thus

$$\begin{aligned}
&\left| \langle w_{i,r_a}, \dot{w}_{j,r_b} \rangle_{L^2} - \left\langle w^{(m)}_{p,a}, \dot{w}^{(m)}_{q,b} \right\rangle_{L^2} \right| \\
&\leq \|w_{i,r_a}\|_{W^{1,1}} \|w_{j,r_b} - w^{(m)}_{q,b}\|_{W^{1,1}} + \|w_{i,r_a} - w^{(m)}_{p,a}\|_{W^{1,1}} \|w^{(m)}_{q,r_b}\|_{W^{1,1}} \\
&\to 0, \quad \text{as } m\to\infty, \text{ by (5.80)}.
\end{aligned}$$

Proving (5.83) next, for p,a as in (5.85) and $u\in\mathscr{U}$ we have

$$(\mathcal{B}^m_p u)_a = \langle \sigma^{(m)}_p v^{(m)}_{p,a}(0), u\rangle_{\mathscr{U}}$$

which is the α^{th}, say, component of the vector obtained by applying the matrix on the right hand side of (5.83) to u. By the convergence of the singular values in (5.29) and $W^{1,1}$ convergence of the Schmidt vectors as in (5.80) we see that

$$\langle \sigma^{(m)}_p v^{(m)}_{p,a}(0), u\rangle_{\mathscr{U}} \to \langle \sigma_i v_{i,r_a}(0), u\rangle_{\mathscr{U}} =: (\mathcal{B}_i u)_\alpha, \quad \text{as} \quad m\to\infty,$$

by construction of $\mathcal{B}_i$. We conclude (5.83) holds. The proof of (5.84) is similar.

The second claim; the convergence in (5.79), follows from (i) as in the proof of [27, Lemma 4.4]. □

Our second main result of this section describes some of the properties of the truncated system.

Proposition 5.3.11. *Let H denote a Hankel operator satisfying assumption* **A** *with* *transfer function G and let G_n denote the transfer function obtained by Lyapunov balanced truncation of G. The realisation $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$ on $(\mathscr{Y}, \mathscr{X}_n, \mathscr{U})$ of G_n from Definition 5.3.5 is stable (that is, A_n is asymptotically stable), minimal and output-normal. Moreover, the Hankel singular values of the Lyapunov balanced truncation are the first n singular values of H, with the same multiplicities.*

The proof of Proposition 5.3.11 is conceptually very similar to that of Pernebo & Silverman [66] for Lyapunov balanced truncation for finite-dimensional systems. A proof for the finite-dimensional case can also be found, for example, in Green & Limebeer [33, Lemma 9.4.1]. The broad idea is to derive some Lyapunov equations that the truncated operators A_n, B_n, C_n and their "adjoints" (in senses we make precise later) satisfy. From here we prove A_n is stable and then the claims that $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$ is minimal and output-normal follows from standard finite-dimensional arguments. Since the operators to be truncated A, B and C are defined on Banach spaces with some inherited Hilbert space structure, we argue carefully and need to collect some technical results beforehand. The proof of Proposition 5.3.11 begins on p. 121.

We make a remark first on the notation we will use from now on. Recall also the interpretation of $\mathscr{X}_n$ from Remark 5.3.7 as a Hilbert space equipped with the L^2 inner product.

Remark 5.3.12. Given a Hankel operator satisfying **A** let A, B, C denote the operators from Lemma 5.3.1, and recall the decompositions and projections of Lemma 5.3.4. We define the decompositions

$$A = \begin{bmatrix} \mathcal{P}_n A|_{\mathscr{X}_n} & \mathcal{P}_n A|_{\mathscr{X}_n^{1,1}} \\ \mathcal{Q}_n A|_{\mathscr{X}_n} & \mathcal{Q}_n A|_{\mathscr{X}_n^{1,1}} \end{bmatrix} =: \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{5.88}$$

$$B = \begin{bmatrix} \mathcal{P}_n B \\ \mathcal{Q}_n B \end{bmatrix} =: \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \tag{5.89}$$

$$C = \begin{bmatrix} C|_{\mathscr{X}_n} & C|_{\mathscr{X}_n^{1,1}} \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{5.90}$$

so that for $n \in \mathbb{N}$ the Lyapunov balanced truncation $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$ from Definition 5.3.5 satisfies $A_n = A_{11}$, $B_n = B_1$ and $C_n = C_1$.

Lemma 5.3.13. *Given the operators and decompositions of Lemmas 5.3.1 and 5.3.4 and the notation of Remark 5.3.12, let ${}^2A_{11}^* : \mathscr{X}_n \to \mathscr{X}_n$ denote the (Hilbert space) adjoint of A_{11} so that*

$$\langle x, A_{11} y \rangle_{L^2} = \langle {}^2A_{11}^* x, y \rangle_{L^2}, \quad \forall\, x, y \in \mathscr{X}_n. \tag{5.91}$$

The operator ${}^2A_{11}^*$ is an extension of

$${}^1A_{11}^* := \mathcal{P}_n A^*|_{\mathscr{X}_n^{1,1} \cap D(A^*)} : \mathscr{X}_n \cap D(A^*) \to \mathscr{X}_n,$$

where A^* is the (adjoint) operator from Lemma 5.3.3. Therefore

$$^1A_{11}^* \subseteq {}^2A_{11}^*,$$

which are equal on $\mathscr{X}_n \cap D(A^*)$, and for simplicity we denote both of these operators by A_{11}^* on $\mathscr{X}_n \cap D(A^*)$. Define also the restrictions

$$\begin{aligned} B_1^* &:= B^*|_{\mathscr{X}_n} : \mathscr{X}_n \to \mathscr{U}\,, \\ B_2^* &:= B^*|_{\mathscr{Z}_n^{1,1}} : \mathscr{Z}_n^{1,1} \to \mathscr{U}\,. \end{aligned}$$

Then the Hilbert space adjoint of B_1^* is $B_1 = \mathcal{P}_n B : \mathscr{U} \to \mathscr{X}_n$ as

$$\langle x, B_1 u\rangle_{L^2} = \langle B_1^* x, u\rangle_{\mathscr{U}}, \quad \forall\, u \in \mathscr{U},\ \forall\, x \in \mathscr{X}_n. \tag{5.92}$$

Proof. For $x, y \in \mathscr{X}_n$

$$\begin{aligned} \langle x, Ay\rangle_{L^2} &= \langle \mathcal{P}_n x, A|_{\mathscr{X}_n} y\rangle_{L^2} = \langle x, \mathcal{P}_n A|_{\mathscr{X}_n} y\rangle_{L^2}, \quad \text{by (5.74)}, \\ &= \langle x, A_{11} y\rangle_{L^2}. \end{aligned} \tag{5.93}$$

If additionally $x \in D(A^*)$ then by the adjoint property (5.71)

$$\begin{aligned} \langle x, Ay\rangle_{L^2} &= \langle A^* x, y\rangle_{L^2} = \langle A^*|_{\mathscr{X}_n \cap D(A^*)} x, \mathcal{P}_n y\rangle_{L^2}, \\ &= \langle \mathcal{P}_n A^*|_{\mathscr{X}_n \cap D(A^*)} x, y\rangle_{L^2}, \quad \text{by (5.74)}, \\ &=: \langle {}^1A_{11}^* x, y\rangle_{L^2}. \end{aligned} \tag{5.94}$$

Comparing (5.93) and (5.94) we obtain

$$\langle x, A_{11} y\rangle_{L^2} = \langle {}^1A_{11}^* x, y\rangle_{L^2}, \quad \forall\, x \in \mathscr{X}_n \cap D(A^*), \quad \forall\, y \in \mathscr{X}_n. \tag{5.95}$$

The Hilbert space adjoint ${}^2A_{11}^*$ satisfies (5.91) by definition, and so the claims of the lemma follow from (5.95) and the unicity of the Hilbert space adjoint.

To prove the claims for B_1^* it suffices to prove (5.92). Let $u \in \mathscr{U}$, $x \in \mathscr{X}_n$ so that

$$\begin{aligned} \langle B_1^* x, u\rangle_{\mathscr{U}} &= \langle B^* x, u\rangle_{\mathscr{U}} = \langle x, Bu\rangle_{\mathscr{U}}, \quad \text{by (5.71)}, \\ &= \langle \mathcal{P}_n x, Bu\rangle_{\mathscr{U}} = \langle x, \mathcal{P}_n Bu\rangle_{\mathscr{U}}, \quad \text{by (5.74)}, \\ &= \langle x, B_1 u\rangle_{\mathscr{U}}, \end{aligned}$$

and so the result follows by the unicity of the Hilbert space adjoint of B_1^*. □

Definition 5.3.14. Given the operators of Lemma 5.3.1 and decompositions of Lemma

5.3.4, recall the operator A_{12} from Remark 5.3.12, given by

$$A_{12} = \mathcal{P}_n A|_{\mathscr{Z}_n^{1,1}} : \mathscr{Z}_n^{1,1} \to \mathscr{X}_n.$$

We denote by A_{12}^* the operator

$$\mathcal{Q}_n A^*|_{\mathscr{X}_n \cap D(A^*)}.$$

Remark 5.3.15. It can be proven that A_{12}^* from Definition 5.3.14 satisfies

$$\langle x, A_{12}y\rangle_{L^2} = \langle A_{12}^* x, y\rangle_{L^2}, \quad \forall\, x \in \mathscr{X}_n \cap D(A^*),\ \forall\, y \in D(A), \tag{5.96}$$

which explains the notation. We do not need this fact for our argument and so omit the proof.

In the next lemma we collect several Lyapunov equations which the operators HH^*, A, B, C, A^* and B^* and their truncations satisfy.

Lemma 5.3.16. *Let H denote a Hankel operator satisfying $\mathbf{A}$ and let $L := HH^*$, which recall is a compact operator*

$$\begin{aligned} &L^1(\mathbb{R}^+; \mathscr{Y}) \to L^1(\mathbb{R}^+; \mathscr{Y}),\\ &L^2(\mathbb{R}^+; \mathscr{Y}) \to L^2(\mathbb{R}^+; \mathscr{Y}),\\ &W^{1,1}(\mathbb{R}^+; \mathscr{Y}) \to W^{1,1}(\mathbb{R}^+; \mathscr{Y}). \end{aligned}$$

Then L satisfies

$$\langle x, Ly\rangle_{L^2} = \langle Lx, y\rangle_{L^2}, \quad \forall\, x \in W^{1,1}(\mathbb{R}^+; \mathscr{Y}),\ \forall\, y \in L^1(\mathbb{R}^+; \mathscr{Y}), \tag{5.97}$$

and

$$\mathcal{Q}_n L|_{\mathscr{X}_n} = 0, \tag{5.98}$$

$$\mathcal{P}_n L|_{\mathscr{Z}_n^1} = 0. \tag{5.99}$$

Define the decomposition

$$L = \begin{bmatrix} \mathcal{P}_n L|_{\mathscr{X}_n} & \mathcal{P}_n L|_{\mathscr{Z}_n^1} \\ \mathcal{Q}_n L|_{\mathscr{X}_n} & \mathcal{Q}_n L|_{\mathscr{Z}_n^1} \end{bmatrix} =: \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix}. \tag{5.100}$$

Let A, B, C, A^*, B^* denote the operators from Lemmas 5.3.1 and 5.3.3. Then the fol-

lowing equations hold on $D(A^*)$

$$A^* + A = 0, \tag{5.101}$$
$$AL + LA^* + BB^* = 0, \tag{5.102}$$

and we have their related inner-product versions

$$\langle Av, w\rangle_{L^2} + \langle v, Aw\rangle_{L^2} + \langle Cv, Cw\rangle_{\mathscr{Y}} = 0, \quad \forall\, v, w \in D(A), \tag{5.103}$$
$$\langle ALv, w\rangle_{L^2} + \langle v, ALw\rangle_{L^2} + \langle B^*v, B^*w\rangle_{\mathscr{Y}} = 0, \quad \forall\, v, w \in \mathscr{X}_n. \tag{5.104}$$

The L^2 inner products in the above two equations are understood as the duality pairing of L^1 and L^∞ (the latter containing $W^{1,1}$). The following truncated equations hold

$$\langle A_{11}x, y\rangle_{L^2} + \langle x, A_{11}y\rangle_{L^2} + \langle C_1x, C_1y\rangle_{\mathscr{Y}} = 0, \quad \forall\, x, y \in \mathscr{X}_n, \tag{5.105}$$
$$\langle A_{11}L_1x, y\rangle_{L^2} + \langle x, A_{11}L_1y\rangle_{L^2} + \langle B_1^*x, B_1^*y\rangle_{\mathscr{Y}} = 0, \quad \forall\, x, y \in \mathscr{X}_n, \tag{5.106}$$

where A_{11}, B_1, B_1^*, C_1 are the operators from Remark 5.3.12 and Lemma 5.3.13. The following truncated operator equations hold on $\mathscr{X}_n \cap D(A^*)$

$$A_{12}^* + A_{21} = 0, \tag{5.107}$$
$$A_{11}L_1 + L_1A_{11}^* + B_1B_1^* = 0, \tag{5.108}$$
$$A_{21}L_1 + L_2A_{12}^* + B_2B_1^* = 0. \tag{5.109}$$

Moreover, the following truncated operator equations hold on $\mathscr{X}_n$

$$^2A_{11}^* + A_{11} + {}^2C_1^*C_1 = 0, \tag{5.110}$$
$$L_1{}^2A_{11}^* + A_{11}L_1 + B_1B_1^* = 0. \tag{5.111}$$

The above operators are given by Remark 5.3.12, Lemma 5.3.13, Definition 5.3.14 and ${}^2C_1^* : \mathscr{Y} \to \mathscr{X}_n$ is the Hilbert space adjoint of C_1.

Proof. Both sides of (5.97) make sense and are finite as

$$x \in W^{1,1} \quad \Rightarrow \quad Lx \in W^{1,1},$$
$$y \in L^1 \quad \Rightarrow \quad Ly \in L^1,$$

and so both sides of (5.97) are the pairing of an element of $W^{1,1}$ and an element of L^1. To prove (5.97) let $x \in W^{1,1}$ and $y \in L^1$. Then as $W^{1,1}$ is dense in L^1 there exists a sequence $(y_m)_{m\in\mathbb{N}}$ such that

$$y_m \xrightarrow{L^1} y, \quad \text{as} \quad m \to \infty.$$

Then x and y_m are elements of L^2 and as $L : L^2 \to L^2$ is self-adjoint on L^2

$$\begin{aligned}\langle x, Ly\rangle_{L^2} &= \lim_{m\to\infty}\langle x, Ly_m\rangle_{L^2} = \lim_{m\to\infty}\langle Lx, y_m\rangle_{L^2}\\ &= \langle Lx, y\rangle_{L^2},\end{aligned}$$

where we have used the continuity of L on L^1 and of the duality product.

We now prove (5.98) and (5.99). Observe that $\mathscr{X}_n$ is the sum of the eigenspaces of L corresponding to the first n eigenvalues, so is L-invariant and thus (5.98) holds.

To prove (5.99) consider $x \in \mathscr{X}_n$ and $y \in \mathscr{Z}_n^1$, so that $\mathcal{Q}_n y = y$ and thus

$$\begin{aligned}\langle Lx, y\rangle_{L^2} &= \langle L|_{\mathscr{X}_n} x, \mathcal{Q}_n y\rangle_{L^2} = \langle \mathcal{Q}_n L|_{\mathscr{X}_n} x, y\rangle_{L^2}, \quad \text{by (5.74)},\\ &= 0, \quad \text{by (5.98)}.\end{aligned}$$

Furthermore, by the self-adjointness of L in equation (5.97),

$$0 = \langle Lx, y\rangle_{L^2} = \langle x, Ly\rangle_{L^2} = \langle \mathcal{P}_n x, L|_{\mathscr{Z}_n^1} y\rangle_{L^2} = \langle x, \mathcal{P}_n L|_{\mathscr{Z}_n^1} y\rangle_{L^2}, \quad \text{by (5.74)}.$$

Therefore $\mathcal{P}_n L|_{\mathscr{Z}_n^1} y \in \mathscr{X}_n$ and from the above is orthogonal to $\mathscr{X}_n$. We infer that

$$\mathcal{P}_n L|_{\mathscr{Z}_n^1} y = 0, \quad \forall\, y \in \mathscr{Z}_n^1,$$

and so (5.99) holds.

We now prove the Lyapunov equations in order. Equation (5.101) is established trivially given the definition of A^* in Lemma 5.3.3. For the second equation (5.102) let $x \in D(A^*) = W_0^{1,1}$, so that from the derivative formula (5.46) for H^*x and Hx we compute

$$\begin{aligned}(AL + LA^*)x(t) &= \frac{d}{dt}(HH^*x)(t) - H(H^*\dot{x})(t)\\ &= -h(t)(H^*x)(0) - H(\underbrace{\frac{d}{dt}H^*x + H^*\dot{x}}_{=-h^*(\cdot)x(0)})(t)\\ &= -h(t)(H^*x)(0), \quad \text{as } x(0) = 0,\\ &= -(BB^*x)(t).\end{aligned}$$

We now prove (5.103). Let $v, w \in D(A)$ and $t \geq 0$. Then

$$\langle \dot{v}(t), w(t)\rangle_{\mathscr{Y}} + \langle v(t), \dot{w}(t)\rangle_{\mathscr{Y}} = \frac{d}{dt}\langle v(t), w(t)\rangle_{\mathscr{Y}}.$$

Integrating both sides over $\mathbb{R}^+$ and using the Fundamental Theorem of Calculus (combined with $v, w \in W^{1,1}$) gives

$$\langle Av, w\rangle_{L^2} + \langle v, Aw\rangle_{L^2} = -\langle v(0), w(0)\rangle_{\mathscr{Y}} = -\langle Cv, Cw\rangle_{\mathscr{Y}},$$

which we can rearrange to give (5.103). Equation (5.104) is proved by first considering for $w_{i,k}, w_{j,r} \in \mathscr{X}_n$ which are eigenvectors of L

$$\begin{aligned}\langle ALw_{i,k}, w_{j,l}\rangle_{L^2} + \langle w_{i,k}, ALw_{j,l}\rangle_{L^2} &= \langle \sigma_i^2 \dot{w}_{i,k}, w_{j,l}\rangle_{L^2} + \langle w_{i,k}, \sigma_j^2 \dot{w}_{j,l}\rangle_{L^2} \\ &= \sigma_i\sigma_j(\langle v_{i,k}, \dot{v}_{j,l}\rangle_{L^2} + \langle \dot{v}_{i,k}, v_{j,l}\rangle_{L^2}) \qquad (5.112)\end{aligned}$$

where we need to establish the equality (5.112). A calculation shows

$$\begin{aligned}\langle w_{i,k}, \dot{w}_{j,l}\rangle_{L^2} &= \frac{1}{\sigma_j}\langle w_{i,k}, \sigma_j \dot{w}_{j,l}\rangle_{L^2} = \frac{1}{\sigma_j}\langle w_{i,k}, \frac{d}{dt}Hv_{j,l}\rangle_{L^2} \\ &= \frac{1}{\sigma_j}\left([\langle w_{i,k}(t), Hv_{j,l}(t)\rangle_{\mathscr{Y}}]_0^\infty - \langle \dot{w}_{i,k}, Hv_{j,l}\rangle_{L^2}\right) \\ &= \frac{1}{\sigma_j}\left(-\langle w_{i,k}(0), Hv_{j,l}(0)\rangle_{\mathscr{Y}} - \langle H^*\dot{w}_{i,k}, v_{j,l}\rangle_{L^2}\right) \\ &= \frac{1}{\sigma_j}\left(-\langle h^*(\cdot)w_{i,k}(0), v_{j,l}\rangle_{L^2} - \langle H^*\dot{w}_{i,k}, v_{j,l}\rangle_{L^2}\right) \\ &= \frac{1}{\sigma_j}\langle -h^*(\cdot)w_{i,k}(0) - H^*\dot{w}_{i,k}, v_{j,l}\rangle_{L^2} \\ &= \frac{1}{\sigma_j}\langle \frac{d}{dt}H^*w_{i,k}, v_{j,l}\rangle_{L^2}, \qquad (5.113)\end{aligned}$$

where we have used in (5.113) the derivative formula for H^*, which is given by (5.46) with h and H replaced by h^* and H^* respectively. Now equation (5.113) becomes

$$\langle w_{i,k}, \dot{w}_{j,l}\rangle_{L^2} = \frac{\sigma_i}{\sigma_j}\langle \dot{v}_{i,k}, v_{j,l}\rangle_{L^2}. \qquad (5.114)$$

A calculation very similar to the one above gives

$$\langle w_{j,l}, \dot{w}_{i,k}\rangle_{L^2} = \frac{\sigma_j}{\sigma_i}\langle \dot{v}_{j,l}, v_{i,k}\rangle_{L^2}$$

which when combined with (5.114) yields (5.112). We proceed to rewrite (5.112) as

$$\begin{aligned}\langle ALw_{i,k}, w_{j,l}\rangle_{L^2} + \langle w_{i,k}, ALw_{j,l}\rangle_{L^2} &= \int_{\mathbb{R}^+} \frac{d}{dt}\langle \sigma_i v_{i,k}(t), \sigma_j v_{j,l}(t)\rangle_{\mathscr{U}}\, dt \\ &= -\langle \sigma_i v_{i,k}(0), \sigma_j v_{j,l}(0)\rangle_{\mathscr{U}}, \\ &= -\langle (H^*w_{i,k})(0), (H^*w_{j,l})(0)\rangle_{\mathscr{U}} \\ &= -\langle B^*w_{i,k}, B^*w_{j,l}\rangle_{\mathscr{U}}. \qquad (5.115)\end{aligned}$$

Equation (5.104) now follows by noting that any $x \in \mathscr{X}_n$ can be expressed as a linear combination of finitely many $w_{i,k}$, and that (5.115) is sesquilinear.

To prove (5.105) we start from (5.103), considered for $x, y \in \mathscr{X}_n \subseteq D(A)$

$$\begin{aligned}
0 &= \langle Ax, y\rangle_{L^2} + \langle x, Ay\rangle_{L^2} + \langle Cx, Cy\rangle_{\mathscr{Y}} \\
&= \langle A|_{\mathscr{X}_n}x, \mathcal{P}_n y\rangle_{L^2} + \langle \mathcal{P}_n x, A|_{\mathscr{X}_n}y\rangle_{L^2} + \langle C|_{\mathscr{X}_n}x, C|_{\mathscr{X}_n}y\rangle_{\mathscr{Y}} \\
&= \langle \mathcal{P}_n A|_{\mathscr{X}_n}x, y\rangle_{L^2} + \langle x, \mathcal{P}_n A|_{\mathscr{X}_n}y\rangle_{L^2} + \langle C|_{\mathscr{X}_n}x, C|_{\mathscr{X}_n}y\rangle_{\mathscr{Y}}, \quad \text{by (5.74)},
\end{aligned}$$

which is (5.105). The proof of (5.106) is similar to that above, starting from (5.104).

To prove (5.107) we apply $\mathcal{Q}_n$ to (5.101), and consider for $x \in D := \mathscr{X}_n \cap D(A^*)$

$$0 = \mathcal{Q}_n(A^* + A)x = \mathcal{Q}_n A^*|_D x + \mathcal{Q}_n A|_{\mathscr{X}_n}x = A_{12}^* x + A_{21}x,$$

where we have used the Definition 5.3.14 for A_{12}^*. Next for $x \in D$ applying $\mathcal{P}_n$ to (5.102) gives

$$0 = \mathcal{P}_n(AL + LA^* + BB^*)x = \mathcal{P}_n AL|_{\mathscr{X}_n}x + \mathcal{P}_n LA^*|_D x + B_1 B_1^* x. \tag{5.116}$$

We consider the first two terms on the right hand side of (5.116) separately. Firstly

$$\begin{aligned}
\mathcal{P}_n LA^*|_D = \mathcal{P}_n L(\mathcal{P}_n + \mathcal{Q}_n)A^*|_D &= \underbrace{\mathcal{P}_n L|_{\mathscr{X}_n}}_{=L_1}\mathcal{P}_n A^*|_D + \underbrace{\mathcal{P}_n L|_{\mathscr{X}_n^{1,1}}}_{=0,\text{ by (5.99)}}\mathcal{Q}_n A^*|_D \\
&= {L_1}^1 A_{11}^* = L_1 A_{11}^*. \qquad (5.117)
\end{aligned}$$

Secondly,

$$\mathcal{P}_n AL|_{\mathscr{X}_n} = \mathcal{P}_n A(P_n + Q_n)L|_{\mathscr{X}_n} = \mathcal{P}_n A|_{\mathscr{X}_n} P_n L|_{\mathscr{X}_n} + \mathcal{P}_n A|_{\mathscr{X}_n^{1,1}} Q_n L|_{\mathscr{X}_n} \tag{5.118}$$

where in (5.118) we have used that L maps $W^{1,1}$ into $W^{1,1}$ so that the compositions

$$P_n L|_{\mathscr{X}_n}, \quad Q_n L|_{\mathscr{X}_n},$$

make sense. Now $P_n L|_{\mathscr{X}_n}$ and L_1 are equal on $\mathscr{X}_n$, as P_n and $\mathcal{P}_n$ are equal on $\mathscr{X}_n = L\mathscr{X}_n$. Additionally, $Q_n L|_{\mathscr{X}_n}$ is equal to $\mathcal{Q}_n L|_{\mathscr{X}_n}$ on $\mathscr{X}_n$, which is the zero map, and so $Q_n L|_{\mathscr{X}_n}$ is also zero. Therefore (5.118) becomes

$$\mathcal{P}_n AL|_{\mathscr{X}_n} = \mathcal{P}_n A|_{\mathscr{X}_n} P_n L|_{\mathscr{X}_n} = A_{11}L_1. \tag{5.119}$$

Combining (5.116), (5.117) and (5.119) gives (5.108).

The proof of (5.109) is very similar to that of (5.108), only instead now we multiply (5.102) by $\mathcal{Q}_n$ instead of $\mathcal{P}_n$. The Lyapunov equations (5.110)-(5.111) follow immediately from the inner product versions (5.105) and (5.106) respectively, where in the second equation we have used the adjoint property of B_1^* in equation (5.92). $\square$

Proof of Proposition 5.3.11: We recap that we need to prove that the system $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$ is stable, minimal and output-normal. These claims will follow in light of (5.110)-(5.111) (where $A_n = A_{11}$, $B_n = B_1$, $C_n = C_1$) by standard arguments once we establish the asymptotic stability of A_{11}. In particular, assuming stability of A_{11}, from (5.111) we see that L_1 is the controllability Gramian of the reduced order system, which with respect to the orthonormal basis $(w_{i,k})_{1\le i\le n}^{1\le k\le p_i}$ for $\mathscr{X}_n$ has matrix representation

$$\operatorname{diag}\{\sigma_1^2 I_{p_1}, \dots, \sigma_n^2 I_{p_n}\}, \quad I_p \text{ identity matrix on } \mathbb{C}^p,$$

which is positive and diagonal. Thus the singular values of the reduced order system are the first n singular values of H.

We therefore concentrate on proving the stability of A_{11}. The argument that A_{11} is stable is based on the argument from Pernebo & Silverman [66] for finite-dimensional Lyapunov balanced truncation.

A short calculation using (5.105) demonstrates that every eigenvalue of A_{11} has non-positive real part. To prove A_{11} is asymptotically stable we argue by contradiction. In light of the previous comment we assume that A_{11} has a purely imaginary eigenvalue λ. Let $Z \subseteq \mathscr{X}_n$ denote the eigenspace of A_{11} corresponding to λ. We observe immediately from (5.105) that for $x \in Z$

$$\begin{aligned} &\langle C_1 x, C_1 x\rangle_{\mathscr{Y}} = -2\mathrm{Re}\,\langle A_{11}x, x\rangle_{L^2} = -2\|x\|_2^2\,(\mathrm{Re}\,\lambda) = 0, \\ &\quad \Rightarrow \quad C_1 x = 0, \end{aligned}$$

or equivalently, C restricted to Z is zero. Since $Cz = z(0)$ we infer that

$$z \in Z \quad \Rightarrow \quad Cz = z(0) = 0, \quad \Rightarrow \quad Z \subseteq D(A^*) = W_0^{1,1}(\mathbb{R}^+; \mathscr{Y}), \tag{5.120}$$

in particular, $Z \subseteq \mathscr{X}_n \cap D(A^*) \neq \{0\}$. Considering (5.105) again for $x \in Z$ and $y \in \mathscr{X}_n$ and using (5.120) we observe

$$\begin{aligned} 0 &= \langle A_{11}x, y\rangle_{L^2} + \langle x, A_{11}y\rangle_{L^2} = \langle \lambda x, y\rangle_{L^2} + \langle {}^2A_{11}^* x, y\rangle_{L^2} \\ &= \langle (\lambda I + {}^2A_{11}^*)x, y\rangle_{L^2}. \end{aligned}$$

As $y \in \mathscr{X}_n$ was arbitrary we conclude that

$${}^2A_{11}^* x = -\lambda x = \overline{\lambda} x, \quad \forall\, x \in Z.$$

Since $Z \subseteq \mathscr{X}_n \cap D(A^*)$, from Lemma 5.3.13 we see that ${}^1A_{11}^*$ and ${}^2A_{11}^*$ are equal on Z

and so

$$A_{11}^* x = {}^1A_{11}^* x = {}^2A_{11}^* x = -\lambda x, \quad \forall\, x \in Z. \tag{5.121}$$

For $x \in Z$, by using the adjoint A_{11}^* property in the Lyapunov equation (5.106) we obtain

$$\langle L_1 x, A_{11}^* x\rangle_{L^2} + \langle A_{11}^* x, L_1 x\rangle_{L^2} + \langle B_1^* x, B_1^* x\rangle_{\mathscr{U}} = 0,$$

which when we rearrange and use (5.121) yields

$$\begin{aligned} \langle B_1^* x, B_1^* x\rangle_{\mathscr{U}} &= -(\langle L_1 x, A_{11}^* x\rangle_{L^2} + \langle A_{11}^* x, L_1 x\rangle_{L^2}) \\ &= -2(\operatorname{Re}\lambda)\langle L_1 x, x\rangle_{L^2} = 0, \\ \Rightarrow \quad B_1^* x &= 0. \end{aligned}$$

We conclude that B^* restricted to Z is zero. Therefore from the truncated equation (5.108) we obtain for $x \in Z$

$$(A_{11} L_1 + L_1 A_{11}^*) x = 0, \tag{5.122}$$

Inserting (5.121) into (5.122) gives

$$A_{11}(L_1 x) = \lambda (L_1 x),$$

and so we infer that Z is L_1-invariant. Now the truncated equation (5.107) yields for $x \in Z$

$$A_{12}^* x = -A_{21} x,$$

which when substituted into (5.109)

$$A_{21} L_1 x + L_2 A_{12}^* x + B_2 \underbrace{B_1^* x}_{=0} = 0,$$

gives

$$A_{21} L_1 = L_2 A_{21}. \tag{5.123}$$

Since Z is L_1-invariant we can restrict L_1 to an operator

$$L_1^r : Z \to Z,$$

and we remark that the spectrum of L_1^r is contained within the spectrum of L_1. So choose an eigenvalue μ of L_1^r, with corresponding eigenvector v. From (5.123) we note that

$$L_2(A_{21} v) = A_{21} L_1 v = \mu (A_{21} v).$$

As L_1^r and L_2 have disjoint spectra, we conclude that $A_{21} v = 0$.

Therefore, the operator A has eigenvector $v \in Z$, corresponding to the eigenvalue λ as

$$Av = A|_{\mathscr{X}_n} v = \mathcal{P}_n A|_{\mathscr{X}_n} v + \mathcal{Q}_n A|_{\mathscr{X}_n} v = A_{11}v + A_{21}v = \lambda v.$$

As such the semigroup τ^t has eigenvector v with eigenvalue $\mathrm{e}^{\lambda t}$. Recall that the output map of ${}^{sr}\Sigma^1$ is the identity, and so using (5.120) we obtain the contradiction

$$v(t) = (Iv)(t) = C\tau^t v = \mathrm{e}^{\lambda t} Cv = 0.$$

We conclude that $A_n = A_{11}$ is asymptotically stable and the rest of the proof follows as outlined on p. 121. □

5.3.3 Relation to earlier work

We remark on some of the connections of the results of this section with [27, Section 3]. As explained in Section 5.2.2, the assumption $h \in L^2$ implies that the Schmidt pairs belong to $W^{1,2}$, which is the domain of the generator of the semigroup of the output-normal realisation ${}^{sr}\Sigma^2$. Therefore the authors of [27] construct an output-normal (L^2 well-posed) realisation of a Hankel operator satisfying their assumptions. The balanced truncation in [27] is obtained from this realisation.

Since we have removed the assumption $h \in L^2$, by Lemma 5.2.12 the Schmidt pairs do *not* belong to $W^{1,2}$. Instead, by Lemma 5.2.10, $h \in L^1$ implies the Schmidt pairs belong to $W^{1,1}$. As we seek realisations where we can naturally describe the truncations in terms of the Schmidt vectors we are forced to consider a well-posed realisation of H on L^1. The exactly observable shift realisation ${}^{sr}\Sigma^1$, which we truncate, is the natural Banach space equivalent of the Hilbert space output-normal realisation.

And, as it turns out, our definition of reduced order system obtained by Lyapunov balanced truncation agrees with that of Glover *et al.* [27, Section 4], in the sense that they define their truncation in terms of the matrices given by (5.78) (once adjusted for multiplicities of the singular values). Proposition 5.3.11 demonstrates that the Lyapunov balanced truncation is stable, output-normal and minimal. Moreover, from Theorem 5.0.2 we see that using this truncation method we obtain the infinite-dimensional version of the Lyapunov balanced truncation error bound (2.7).

The conclusion (*ii*) of Proposition 5.3.9 is the same as one of the conclusions of [27, Lemma 4.4], but crucially does not use the additional assumptions in [27]. Specifically, if $h \in L^2$ then the sequence $(h_m)_{m\in\mathbb{N}}$ chosen so that

$$h_m \xrightarrow{L^1, L^2} h, \quad \text{as} \quad m \to \infty,$$

(see Remark 5.2.15) is used in [27, Lemma 4.3, (iii)] to prove convergence of the Schmidt pairs in L^∞ and so also at zero. Convergence at zero gives component wise convergence of the matrices $\mathcal{B}_i^m$, $\mathcal{C}_i^m$. The assumption that h is real or $\dot{h}$ exists and is the kernel of a bounded Hankel operator is used in [27, Lemma 4.4] to prove component wise convergence of the $\mathcal{A}_{ij}^m$. This assumption is unnecessary, and is avoided by establishing $W^{1,1}$ convergence of the Schmidt pairs in Theorem 5.2.2.

Throughout Section 5.3 we did not need to assume that H is nuclear, only that assumption **A** holds.

5.4 Proof of the Lyapunov balanced truncation error bound

The proof is similar to that of [27, Theorem 5.1], only the technical results of [27] have been replaced with ours to accommodate our weaker assumptions. Specifically [27, Lemma 4.4] has been replaced by Proposition 5.3.9. We have also taken into consideration the multiplicities of the singular values.

Proof of Theorem 5.0.2: We apply Proposition 5.3.9 to the partial sums of the Coifman & Rochberg decompositions from Corollary 5.1.14 and Lemma 5.1.17. That is, for $m \in \mathbb{N}$ define

$$h_m(t) := \sum_{j=1}^{m} \lambda_j (\operatorname{Re} a_j) \mathrm{e}^{a_j t}, \quad t > 0, \quad G^m(s) := \sum_{j=1}^{m} \lambda_j \frac{\operatorname{Re} a_j}{s - a_j}, \quad \operatorname{Re} s > 0, \tag{5.124}$$

so that by Corollary 5.1.14 the sequence $(H_m)_{m\in\mathbb{N}}$ given by (5.28) converges in nuclear norm to H. The following chain of inequalities holds

$$\|G - G^m\|_{H^\infty} \leq \|h - h_m\|_1 \leq 2\|H - H_m\|_N, \tag{5.125}$$

where the second is proven in [27, Theorem 2.1]. Therefore the impulse responses h_m converge to h in L^1 and so the conditions of Proposition 5.3.9 are satisfied. We first prove that

$$\|G - G_n\|_{H^\infty} \to 0, \quad \text{as} \quad n \to \infty. \tag{5.126}$$

Let $\varepsilon > 0$ be given. For $n, m \in \mathbb{N}$, the triangle inequality yields

$$\|G - G_n\|_{H^\infty} \leq \|G - G^m\|_{H^\infty} + \|G^m - G_n^m\|_{H^\infty} + \|G_n^m - G_n\|_{H^\infty}. \tag{5.127}$$

Our aim is to show that each of the summands on the right-hand side of (5.127) can be made smaller than (a fraction of) ε. Choose $M_1 \in \mathbb{N}$ such that for $m \geq M_1$

$$\|H - H^m\|_N < \frac{\varepsilon}{12}, \tag{5.128}$$

so that by (5.125) for $m \geq M_1$

$$\|G - G^m\|_{H^\infty} < \frac{\varepsilon}{6}. \tag{5.129}$$

Secondly, by Proposition 5.3.9 for each $n \in \mathbb{N}$ there exists $M_2 \in \mathbb{N}$ (which depends on n) such that $m \in \mathbb{N}$ and $\tau(m) \geq m \geq M_2(n)$ implies that

$$\|G_n^{(\tau(m))} - G_n\|_{H^\infty} < \frac{\varepsilon}{3}, \tag{5.130}$$

where $\tau(m)$ denotes the m^{th} term of the subsequence from Proposition 5.3.9. Recall the sequence of increasing integers l_i from Theorem 5.2.2. For each $n \in \mathbb{N}$ let $q \in \mathbb{N}_0$ be such that

$$l_q \leq n < l_{q+1}. \tag{5.131}$$

Choose $N \in \mathbb{N}$ such that $n \geq N$ implies that

$$\sum_{k=q+1}^{\infty} p_k \sigma_k < \frac{\varepsilon}{12}, \tag{5.132}$$

which is possible by the nuclearity of H and the choice of q in (5.131). Now for $n \geq N$ choose M_3 (which depends on n) such that $m \geq M_3(n)$ implies that

$$\left| \sum_{j=1}^{l_q} p_j^{(m)} \sigma_j^{(m)} - \sum_{j=1}^{q} p_j \sigma_j \right| < \frac{\varepsilon}{12}, \tag{5.133}$$

which is possible by our choice of l_q in (5.131) and the convergence in (5.29). Note that as G_m is rational for each $m \in \mathbb{N}$, its Hankel operator H^m is finite rank. Therefore for each $m \in \mathbb{N}$ the sequence of singular values $(\sigma_j^{(m)})_{j \in \mathbb{N}}$ contains only finitely many non-zero terms. We let $\mathcal{N}(m)$ denote the number of non-zero (and therefore distinct) singular values of H^m. For $n \geq N$ choose $m \in \mathbb{N}$ such that $m \geq \max\{M_1, M_2, M_3\}$ and $\mathcal{N}(m) \geq n + 1$. The Lyapunov balanced truncation error bound for rational transfer functions, Theorem 2.1.9, applies to second term of (5.127) to give

$$\|G^m - G_n^m\|_{H^\infty} \leq 2 \sum_{k=n+1}^{\mathcal{N}(m)} \sigma_k^{(m)} \leq 2 \sum_{k=n+1}^{\infty} \sigma_k^{(m)}. \tag{5.134}$$

We proceed to show that for our choice of n and m the right-hand side of (5.134) is (arbitrarily) small. We have

$$\begin{aligned}
2\sum_{k=n+1}^{\infty}\sigma_k^{(m)} &\le 2\sum_{k=n+1}^{\infty}p_k^{(m)}\sigma_k^{(m)} \le 2\sum_{k=l_q+1}^{\infty}p_k^{(m)}\sigma_k^{(m)} \\
&= 2\left(\sum_{k=l_q+1}^{\infty}p_k^{(m)}\sigma_k^{(m)} - \sum_{k=q+1}^{\infty}p_k\sigma_k\right) + 2\sum_{k=q+1}^{\infty}p_k\sigma_k \\
&= 2\left(\sum_{k\in\mathbb{N}}p_k^{(m)}\sigma_k^{(m)} - \sum_{k\in\mathbb{N}}p_k\sigma_k\right) + 2\sum_{k=q+1}^{\infty}p_k\sigma_k \\
&\quad + 2\left(\sum_{k=1}^{l_q}p_k^{(m)}\sigma_k^{(m)} - \sum_{k=1}^{q}p_k\sigma_k\right).
\end{aligned}$$

Therefore, by the triangle inequality

$$\begin{aligned}
2\left|\sum_{k=n+1}^{\infty}\sigma_k^{(m)}\right| &\le 2\left|\sum_{k\in\mathbb{N}}p_k^{(m)}\sigma_k^{(m)} - \sum_{k\in\mathbb{N}}p_k\sigma_k\right| + 2\sum_{k=q+1}^{\infty}p_k\sigma_k \\
&\quad + 2\left|\sum_{k=1}^{l_q}p_k^{(m)}\sigma_k^{(m)} - \sum_{k=1}^{q}p_k\sigma_k\right| \qquad (5.135) \\
&< 2\|H - H_m\|_N + \frac{\varepsilon}{6} + \frac{\varepsilon}{6}, \qquad (5.136)
\end{aligned}$$

where we have bounded the second and third terms in (5.135) by (5.132) and (5.133) respectively. Now using the bound (5.128) in (5.136) and combining with (5.134) we obtain for n, m as above

$$\|G^m - G_n^m\|_{H^\infty} \le 2\left|\sum_{k=n+1}^{\infty}\sigma_k^{(m)}\right| < \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} < \frac{\varepsilon}{3}. \qquad (5.137)$$

Putting the bounds (5.129), (5.130) and (5.137) together now, we see that there exists $N \in \mathbb{N}$ such that for $n \ge N$ and $m \in \mathbb{N}$ such that $\tau(m) \ge m \ge \max\{M_1, M_2, M_3\}$ and $\mathcal{N}(\tau(m)) \ge \mathcal{N}(m) \ge n + 1$ we have

$$\begin{aligned}
\|G - G_n\|_{H^\infty} &\le \|G - G^{(\tau(m))}\|_{H^\infty} + \|G^{(\tau(m))} - G_n^{(\tau(m))}\|_{H^\infty} \\
&\quad + \|G_n^{(\tau(m))} - G_n\|_{H^\infty} \\
&< \frac{\varepsilon}{6} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon,
\end{aligned}$$

proving (5.126). To prove the error bound we use Theorem 2.1.9 again to obtain that for $j > n$

$$\|G_j - G_n\|_{H^\infty} \le 2\sum_{k=n+1}^{j}\sigma_k \le 2\sum_{k=n+1}^{\infty}\sigma_k. \qquad (5.138)$$

For the above we have used that the output-normal realisation $\left[\begin{smallmatrix} A_n & B_n \\ C_n & 0 \end{smallmatrix}\right]$ of G_n is the balanced truncation of the output-normal realisation $\left[\begin{smallmatrix} A_j & B_j \\ C_j & 0 \end{smallmatrix}\right]$ of G_j, which follows from Proposition 5.3.11. To obtain the error bound (5.7), let $\varepsilon > 0$ be given and so by (5.126) we can choose $j \in \mathbb{N}$, $j > n$ such that

$$\|G - G_n\|_{H^\infty} \leq \|G - G_j\|_{H^\infty} + \|G_j - G_n\|_{H^\infty} \leq \varepsilon + 2 \sum_{k=n+1}^{\infty} \sigma_k,$$

where we have used (5.138) to bound the second term above. Since $\varepsilon > 0$ was arbitrary, we conclude that the error bound holds. □

Remark 5.4.1. Note the σ_k are *not* repeated in the error bound according to multiplicity (which is also the case for the finite-dimensional bound).

5.5 Applications of Lyapunov balanced truncation

5.5.1 Optimal Hankel-norm approximations

We comment briefly on the optimal Hankel norm approximations of Glover *et al.* [27, Section 6]. The following result is based on [27, Theorem 6.4] and demonstrates how the key assumption is nuclearity of the Hankel operator. We refer the reader to that article for the full details.

Theorem 5.5.1. *Suppose G is the transfer function corresponding to a nuclear Hankel operator with singular values $(\sigma_n)_{n\in\mathbb{N}}$ and fix an integer k. Then there exists a transfer function $\hat{G}^\infty$ of MacMillan degree k such that*

$$\|G - \hat{G}^\infty\|_H = \sigma_{k+1},$$

where $\|F\|_H$ denotes the Hankel norm of the Hankel operator corresponding to the transfer function F. Thus $\hat{G}^\infty$ is an optimal-Hankel norm approximant for G. Moreover, there exists a constant matrix D_0 such that

$$\|G - \hat{G}^\infty - D_0\|_{H^\infty} \leq \sum_{n=k+1}^{\infty} \sigma_n.$$

Proof. The proof is the same as in [27, Theorem 6.4], only with [27, Theorem 5.1] replaced by Theorem 5.0.2. □

5.5.2 Application to numerical algorithms

Theorem 5.0.3, with a very particular choice of the sequence h_m, is used in the proof of Theorem 5.0.2. However it is also of independent interest in connection to numerical algorithms for approximating balanced truncations.

In this case the sequence h_m is obtained as the impulse response of a finite-dimensional state space system (A_m, B_m, C_m). Usually h is the impulse response of a controlled partial differential equation and (A_m, B_m, C_m) is obtained from a semi-discretization of this partial differential equation. The convergence $h_m \xrightarrow{L^1} h$ assumed in Theorem 5.0.3 is typical in such situations.

A different approach to model reduction using approximations of Schmidt vectors of Hankel operators is taken by [74]. We recall some key ideas of that approach here to compare it to the situation of Theorem 5.0.3.

The starting point in [74] is a linear system with generators (A, B, C). It is assumed that $A : X \supset D(A) \to X$ generates an exponentially stable semigroup and X is an (infinite dimensional) Hilbert space with (real) inner product $\langle \cdot, \cdot \rangle_X$. The input and output spaces are $\mathbb{R}^m$ and $\mathbb{R}^p$ respectively. Let $\{e_1, \dots, e_m\}$ and $\{e_1, \dots, e_p\}$ denote the standard orthonormal bases for $\mathbb{R}^m$ and $\mathbb{R}^p$ respectively. The operators B and C are assumed finite rank and bounded, and so can be formulated

$$Bu = \sum_{k=1}^{m} b_k u_k, \quad \text{where} \quad u = \sum_{k=1}^{m} u_k e_k = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}, u_k \in \mathbb{R},$$

$$Cx = \sum_{l=1}^{p} \langle c_l, x \rangle_X e_l.$$

Define z_i and w_j as the solutions of the initial value problems

$$\dot{z}_i(t) = A^* z_i(t), \quad z_i(0) = c_i \in X, \tag{5.139}$$

$$\dot{w}_j(t) = A w_j(t), \quad w_j(0) = b_j \in X, \tag{5.140}$$

where $i \in \{1, 2, \dots, p\}$ and $j \in \{1, 2, \dots, m\}$.

Under the above assumptions the Hankel operator H of the system generated by (A, B, C) is nuclear and is given by the integral operator (5.2) with kernel $h \in L^1$. With respect to the above bases for $\mathbb{R}^m$ and $\mathbb{R}^p$, h has the matrix representation

$$h_{ij}(t+s) = \langle e_i, h(t+s) e_j \rangle_{\mathbb{R}^p} = \langle z_i(t), w_j(s) \rangle_X, \quad \forall\, t, s \geq 0, \tag{5.141}$$

and all $i \in \{1, 2, \dots, p\}$ and $j \in \{1, 2, \dots, m\}$.

A sequence of approximations is obtained by approximating the data z_i, w_j by z_i^N, w_j^N in $L^2(\mathbb{R}^+; X)$ and defining k_N by

$$(k_N)_{ij}(t, s) := \langle z_i^N(t), w_j^N(s) \rangle_X, \quad \forall\, t, s \geq 0. \tag{5.142}$$

If the approximations z_i^N and w_j^N are obtained by a semi-discretization, then in fact $k_N(t,s) = h_N(t+s)$ with h_N the impulse response of the semi-discretization. In this case, the assumption that

$$\begin{aligned} z_i^N &\to z_i, \\ w_j^N &\to w_j, \end{aligned} \quad \text{in } L^2(\mathbb{R}^+; X) \text{ for all } i, j \text{ as } N \to \infty, \tag{5.143}$$

together with equations (5.141) and (5.142) implies that

$$h_N \xrightarrow{L^1} h, \quad \text{as} \quad N \to \infty.$$

Therefore the assumptions (5.143) from [74] imply that the hypotheses of Theorem 5.0.3 hold in case that the approximations are obtained by semi-discretization (the latter is however not the approach suggested in [74] to approximate the solutions of (5.139) and (5.140). Moreover, in [74] instead also a discretization in time is suggested).

Another difference is that in [74] no balanced truncation of h_N is performed. Instead a snapshot and a quadrature approach are suggested to approximate the Schmidt pairs of the Hankel operator H. Those approximations are then used to define a reduced order system. This alternative approach necessitates making stronger convergence assumptions than (5.143). However, these alternative approaches avoid having to compute balanced truncations of a large-scale system and are therefore computationally more attractive.

5.6 Notes

The contents of this chapter have been submitted for publication as [36]. In that article the model reduction process was called balanced truncation, as opposed to Lyapunov balanced truncation, as we have not truncated an output-normal, let alone a Lyapunov balanced, realisation. We have subsequently changed the naming convention for this thesis for two reasons. Firstly, because the process gives rise to the same reduced transfer function and secondly, so as to distinguish it from bounded real and positive real balanced truncation, described in Chapters 6 and 7 respectively.

As we mentioned in the introduction and the start of this chapter, the results presented here can largely be viewed as an extension of [27]. These extensions were necessary because in deriving (our primary goal of) bounded real and positive real balanced truncation we required the results of that article, but their assumptions were too restrictive. We also feel the results are of independent interest. We have tried to highlight within the chapter where the differences and novelties arise, see especially Sections 5.2.2 and 5.3.3. The only assumption in Theorem 5.0.2 is that the Hankel operator is nuclear. The question of which systems have nuclear Hankel operators has been addressed in, for example, Curtain & Sasane [19] and Opmeer [61].

The material on Hankel operators is known, but we had difficulty finding it in the literature precisely in the form presented here, hence its inclusion. The material of that section was based largely on [54], [62], [63] and [65].

As mentioned in the introduction, the existence of balanced realisations of non-rational transfer functions is non-trivial. Existence of balanced realisations was proven in the discrete time case by Young [103] and converted to general continuous-time systems by Ober & Montgomery-Smith [56]. For systems satisfying the assumptions of [27], balanced and output-normal realisations are considered in Curtain & Glover [17] and [27, Section 3] respectively. More recently, output-normal and balanced realisations have been described for L^2 well-posed linear systems in Staffans [81, Chapter 9].

Chapter 6

Bounded real balanced truncation

In this chapter we extend bounded real balanced truncation to a class of infinite-dimensional systems. The main result of this chapter is Theorem 6.3.15 which states that for strictly bounded real G with summable bounded real singular values $(\sigma_k)_{k\in\mathbb{N}}$ for each $n \in \mathbb{N}$ there exists a rational bounded real transfer function $G_n \in H^\infty$ such that the error bound

$$\|G - G_n\|_\infty \leq 2 \sum_{k=n+1}^{N} \sigma_k,$$

holds. Many infinite-dimensional systems, such as controlled time dependent partial differential equations, incorporate energy dissipation and hence are often either positive real or bounded real. There are many model reduction schemes that can be employed to compute numerical solutions of such PDEs, for example the spatial discretisation finite-element or finite-difference methods. These methods give rise to finite-dimensional approximations of the original system. If a finite-element approximation is based on the physically motivated energy norm (usually a conservation law) of the original PDE, also energy notions (i.e. bounded realness or positive realness of the transfer function) are approximated correctly.

However, approximation of controlled partial differential equations by such numerical methods often gives results that are far from optimal [58]. A rigorous verification of this observation depends on two things: 1) an error analysis of these standard numerical methods and 2) determining what the optimal approximation results (approximately) are. The above error bound, combined with the trivial lower bound

$$\sigma_{n+1} \leq \|G - G_n\|_\infty,$$

which holds for any reduced order system of dimension n demonstrates that the bounded real balanced truncation is indeed close to optimal. In Chapter 8 we provide an analysis of the bounded real singular values which shows that in many of the above applications these singular values converge to zero at a rate faster than any polynomial rate (whether the rate is in fact exponential is –for partial differential equation examples– an open problem) [58, 61]. This implies that bounded real balanced truncations in these applications converge at a very fast rate. Standard numerical methods such as finite-elements do not converge as quickly in these applications which motivates bounded real balanced truncation as an effective model reduction scheme. We give an example highlighting these varying convergence rates in Chapter 8.

6.1 Approach for the infinite-dimensional case

Existence of bounded real balanced realisations in the infinite-dimensional case is shown in [81, Theorem 11.8.14], however bounded real balanced truncation is not addressed there. Deriving bounded real balanced truncation in the infinite-dimensional case by repeating the entire construction in Section 2.2 is technically much more involved. This is, loosely speaking, because the Bounded Real Lemma doesn't exist as concisely in the infinite-dimensional case. Results in that direction do exist in for example, [97] and Arov & Staffans [5], but still it is harder to write down the relationship between the optimal cost operators P_m and P_M (the latter of which is unbounded in interesting cases) and the extremal solutions of (2.10). Furthermore, and unlike the finite-dimensional case, here the difference between non-strict and strict is much greater. For instance, the example of Weiss & Zwart [96] demonstrates that even in the strictly bounded real case the bounded real algebraic Riccati equation (2.11) does not hold in its current form.

Therefore, in the infinite-dimensional case we construct the bounded real balanced truncation by relating it to the Lyapunov balanced truncation of a certain extended system, as outlined (for the finite-dimensional case) in Section 2.2.3. To that end we draw on the material developed in Chapter 5 on Lyapunov balanced truncation.

Many of the results in (finite-dimensional) positive real balanced truncation follow from the corresponding results in bounded real balanced truncation and vice versa via the Cayley transform. Although historically positive real balanced truncation was derived first, it seems more natural for us in infinite-dimensions to consider the bounded real case first and then treat the positive real case using the Cayley transform. There are two reasons for this. Firstly, because bounded real systems are more general (they need not be square, i.e. $\mathscr{U} \neq \mathscr{Y}$ is permitted) and secondly, because of the connections we've described to (the known) Lyapunov balanced truncation, which as described in Remark 2.3.7, do not hold for positive real balanced truncation.

6.2 Extended systems

Our starting point is a strictly bounded real transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, where $\mathscr{U}$ and $\mathscr{Y}$ are the input and output spaces respectively, which as always are assumed to be finite-dimensional Hilbert spaces. In contrast to Section 2.2, we are not assuming that G is rational and at present we make no regularity assumptions on G. We are required to assume to that G is strictly bounded real, in the first instance in order to apply the optimal control results [97]. We assume that $\mathscr{U}$ and $\mathscr{Y}$ are finite-dimensional so that the Lyapunov balanced truncation results of Chapter 5 apply. See in particular Remark 5.1.16.

For bounded real balanced truncation we need a state-space realisation of G. Since $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, by Lemma 4.1.5 there exist stable L^2 well-posed realisations of G. Bounded real balanced truncation makes use of the unique optimal cost operator of the optimal control problem below. The following result is taken from [97], but can also be found in Staffans [77].

Lemma 6.2.1. *Let $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ denote a stable L^2 well-posed linear system and assume that Σ has strictly bounded real transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$. Then the optimal control problem: for $x_0 \in \mathscr{X}$ minimise*

$$\mathcal{J}(x_0, u) = \int_{\mathbb{R}^+} \|u(s)\|_{\mathscr{U}}^2 - \|y(s)\|_{\mathscr{Y}}^2 \, ds, \tag{6.1}$$

over all $u \in L^2(\mathbb{R}^+; \mathscr{U})$ subject to (4.1)*, has a solution in the sense that for any $x_0 \in \mathscr{X}$*

$$\inf_{u \in L^2(\mathbb{R}^+; \mathscr{U})} \mathcal{J}(x_0, u) = \mathcal{J}(x_0, u_{opt}) = -\langle P_m x_0, x_0 \rangle_{\mathscr{X}}. \tag{6.2}$$

The optimal control is uniquely given by

$$u_{opt} = (I - \pi_+ \mathfrak{D}^* \mathfrak{D} \pi_+)^{-1} \pi_+ \mathfrak{D}^* \mathfrak{C} x_0, \tag{6.3}$$

and $P_m : \mathscr{X} \to \mathscr{X}$ is bounded and satisfies $P_m = P_m^* \geq 0$ and

$$P_m = \mathfrak{C}^* \mathfrak{C} + \mathfrak{C}^* \mathfrak{D} \pi_+ (I - \pi_+ \mathfrak{D}^* \mathfrak{D} \pi_+)^{-1} \pi_+ \mathfrak{D}^* \mathfrak{C}. \tag{6.4}$$

Proof. See [97, Proposition 7.2]. Note that the assumption that G is strictly bounded real is equivalent to

$$I - \pi_+ \mathfrak{D}^* \mathfrak{D} \pi_+ \geq \varepsilon I,$$

see [97, Section 7] and hence $I - \pi_+ \mathfrak{D}^* \mathfrak{D} \pi_+$ is boundedly invertible. Therefore the optimal control u_{opt} and optimal cost operator P_m in (6.3) and (6.4) respectively are well-defined. Furthermore, in [97] it is assumed that G is weakly regular (with zero feedthrough), but that is not needed for this proof. □

The dual optimal control problem is now formulated and solved as before. It is easy to see from Definition 4.2.1 of the dual transfer function that $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is (strictly) bounded real if and only if G_d is.

Lemma 6.2.2. *Given a stable L^2 well-posed linear system Σ with strictly bounded real transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$, let Σ_d denote the dual system from Definition 4.2.1. Then the dual optimal control problem: for each $x_0 \in \mathscr{X}$ minimise*

$$\mathcal{J}_d(x_0, y) = \int_{\mathbb{R}^+} \|y_d(s)\|_{\mathscr{Y}}^2 - \|u_d(s)\|_{\mathscr{U}}^2 \, ds, \tag{6.5}$$

over all $y_d \in L^2(\mathbb{R}^+; \mathscr{Y})$ subject to (4.9), has a solution in the sense that for any $x_0 \in \mathscr{X}$

$$\inf_{y_d \in L^2(\mathbb{R}^+;\mathscr{Y})} \mathcal{J}_d(x_0, y_d) = \mathcal{J}_d(x_0, y_{d,opt}) = -\langle Q_m x_0, x_0 \rangle_{\mathscr{X}}. \tag{6.6}$$

The optimal control is uniquely given by

$$y_{d,opt} = (I - \pi_+({}^d\mathfrak{D})^*({}^d\mathfrak{D})\pi_+)^{-1}\pi_+({}^d\mathfrak{D})^{*\,d}\mathfrak{C}x_0, \tag{6.7}$$

and $Q_m : \mathscr{X} \to \mathscr{X}$ is bounded and satisfies $Q_m = Q_m^* \geq 0$ and

$$Q_m := ({}^d\mathfrak{C})^{*\,d}\mathfrak{C} + ({}^d\mathfrak{C})^{*\,d}\mathfrak{D}\pi_+(I - \pi_+({}^d\mathfrak{D})^{*\,d}\mathfrak{D}\pi_+)^{-1}\pi_+({}^d\mathfrak{D})^{*\,d}\mathfrak{C}. \tag{6.8}$$

We seek to extend the original transfer function G to a "larger" system with Hankel operator that has the same singular values as the product $Q_m P_m$. To that end we draw on the results on spectral factorisations and particularly spectral factor systems developed in [97]. This is the second instance of where we require strict bounded realness of G.

We proceed in stages; firstly in Lemmas 6.2.4 and 6.2.5 below we construct two families of intermediate extended systems which we combine in Lemma 6.3.7 and Definition 6.3.8.

Lemma 6.2.3. *If $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is a strictly bounded real transfer function, then there exist functions θ satisfying $\theta, \theta^{-1} \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ and ξ satisfying $\xi, \xi^{-1} \in H^\infty(\mathbb{C}_0^+; B(\mathscr{Y}))$ such that*

$$I - [G(\mathrm{i}\omega)]^* G(\mathrm{i}\omega) = [\theta(\mathrm{i}\omega)]^* \theta(\mathrm{i}\omega), \quad \text{for almost all } \omega \in \mathbb{R}, \tag{6.9}$$

and

$$I - G(\mathrm{i}\omega)[G(\mathrm{i}\omega)]^* = \xi(\mathrm{i}\omega)[\xi(\mathrm{i}\omega)]^*, \quad \text{for almost all } \omega \in \mathbb{R}. \tag{6.10}$$

The functions θ and ξ are uniquely determined up to multiplication by a unitary operator *in $B(\mathscr{U})$ and $B(\mathscr{Y})$ respectively. Specifically, if θ_0 satisfies* (6.9) *and ξ_0 satisfies* (6.10) *then the sets of all spectral factors satisfying* (6.9) *and* (6.10) *are given by*

$$\{U\theta_0 \,:\, U \in B(\mathscr{U}),\, U \text{ unitary}\} \quad \text{and} \quad \{\xi_0 V \,:\, V \in B(\mathscr{Y}),\, V \text{ unitary}\}, \tag{6.11}$$

respectively.

Proof. The assumption that G is strictly bounded real implies that

$$I - [G(\mathrm{i}\omega)]^* G(\mathrm{i}\omega) \geq \varepsilon I, \quad \text{for almost all } \omega \in \mathbb{R}.$$

The existence of the spectral factor θ satisfying $\theta, \theta^{-1} \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, the equality (6.9) and unique up to unitary transformation follows from Rosenblum & Rovnyak [71, Theorem 3.7]. The claims regarding ξ follow from the above and duality. □

Lemma 6.2.4. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ a stable L^2 well-posed realisation. Let $\theta \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a spectral factor from Lemma 6.2.3 satisfying* (6.9) *with input-output map $\mathfrak{D}_\theta$. Define*

$$\mathfrak{C}_E := \begin{bmatrix} \mathfrak{C} \\ \mathfrak{C}_\theta \end{bmatrix} : \mathscr{X} \to L^2\left(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]\right), \tag{6.12}$$

$$\mathfrak{D}_{E_1} := \begin{bmatrix} \mathfrak{D} \\ \mathfrak{D}_\theta \end{bmatrix} : L^2(\mathbb{R}; \mathscr{U}) \to L^2\left(\mathbb{R}; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]\right), \tag{6.13}$$

where

$$\mathfrak{C}_\theta := -\pi_+ \mathfrak{D}_\theta^{-*} \mathfrak{D}^* \mathfrak{C} : \mathscr{X} \to L^2(\mathbb{R}^+; \mathscr{U}). \tag{6.14}$$

In the above $\mathfrak{D}_\theta^{-*} = (\mathfrak{D}_\theta^{-1})^*$. Then $\mathfrak{C}_E$ is bounded and $\Sigma_{E_1} := (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}_E, \mathfrak{D}_{E_1})$ is a stable L^2 well-posed linear system on $(\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right], \mathscr{X}, \mathscr{U})$, with transfer function

$$G_{E_1} := \begin{bmatrix} G \\ \theta \end{bmatrix} \in H^\infty\left(\mathbb{C}_0^+; B\left(\mathscr{U}, \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix}\right)\right), \tag{6.15}$$

and observability Gramian P_m given by (6.2)*, i.e. the optimal cost operator of the optimal control problem* (6.1)*.*

Proof. By [81, Theorem 10.3.5] (alternatively [94, Theorem 1.3]), to the H^∞ function θ we can associate a time invariant, causal, bounded operator

$$\mathfrak{D}_\theta : L^2(\mathbb{R}; \mathscr{U}) \to L^2(\mathbb{R}; \mathscr{U}).$$

The operator $\mathfrak{D}_\theta$ is boundedly invertible since θ^{-1} exists and $\theta^{-1} \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$,

$\mathfrak{D}_\theta^{-1}$ is causal and

$$I - \mathfrak{D}^*\mathfrak{D} = \mathfrak{D}_\theta^*\mathfrak{D}_\theta, \quad \Rightarrow \quad I - \pi_+\mathfrak{D}^*\mathfrak{D}\pi_+ = \pi_+\mathfrak{D}_\theta^*\mathfrak{D}_\theta\pi_+, \tag{6.16}$$

which follows from [97, Section 11] (see particularly (11.5) in the numbering of [97]).

The arguments that follow are based on [97, Theorem 11.1] and [97, Theorem 11.3], only adjusted for our notation. For convenience, we give the arguments. Firstly, for a time invariant, causal bounded operator

$$\tilde{\mathfrak{D}} : L^2(\mathbb{R};\mathscr{U}) \to L^2(\mathbb{R};\mathscr{U}),$$

we claim that

$$\mathfrak{C}_{\text{new}} = -\pi_+\tilde{\mathfrak{D}}^*\mathfrak{C},$$

is an output map for $\mathfrak{A}$ (equivalently, an extended output map in the language of [97]). To see this we need to check that for every $x \in \mathscr{X}$ and all $t \geq 0$

$$\mathfrak{C}_{\text{new}}\mathfrak{A}^t x = \pi_+\tau^t\mathfrak{C}_{\text{new}}x. \tag{6.17}$$

We have

$$\begin{aligned}\mathfrak{C}_{\text{new}}\mathfrak{A}^t x &= -\pi_+\tilde{\mathfrak{D}}^*\mathfrak{C}\mathfrak{A}^t x = -\pi_+\tilde{\mathfrak{D}}^*\pi_+\tau^t\mathfrak{C}x = -\pi_+\tilde{\mathfrak{D}}^*(I-\pi_-)\tau^t\mathfrak{C}x \\ &= -\pi_+\tilde{\mathfrak{D}}^*\tau^t\mathfrak{C}x,\end{aligned}$$

where we have used that $\mathfrak{C}$ is an output map for $\mathfrak{A}$ and also that $\tilde{\mathfrak{D}}^*$ is anticausal, i.e. $\pi_+\tilde{\mathfrak{D}}^*\pi_- = (\pi_-\tilde{\mathfrak{D}}\pi_+)^* = 0$. Now using the time invariance of $\tilde{\mathfrak{D}}$ gives

$$\begin{aligned}\mathfrak{C}_{\text{new}}\mathfrak{A}^t x &= -\pi_+\tau^t\tilde{\mathfrak{D}}^*\mathfrak{C}x = -\pi_+\tau^t(\pi_+ + \pi_-)\tilde{\mathfrak{D}}^*\mathfrak{C}x = -\pi_+\tau^t\pi_+\tilde{\mathfrak{D}}^*\mathfrak{C}x \\ &= \pi_+\tau^t\mathfrak{C}_{\text{new}}x, \quad \text{since } \pi_+\tau^t\pi_- = 0, \text{ for } t \geq 0,\end{aligned}$$

as required. Next note that the product of time invariant, causal, bounded operators is a time invariant, causal, bounded operator. Therefore we apply the above result to $\tilde{\mathfrak{D}} := \mathfrak{D}\mathfrak{D}_\theta^{-1}$ to infer that $\mathfrak{C}_\theta$ given by (6.14) is an output map for $\mathfrak{A}$. We now claim that $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C}_\theta, \mathfrak{D}_\theta)$ is an L^2 well-posed realisation of θ on $(\mathscr{U}, \mathscr{X}, \mathscr{U})$. It remains to check condition (iv) in [81, Definition 2.2.1], namely whether

$$\pi_+\mathfrak{D}_\theta\pi_- = \mathfrak{C}_\theta\mathfrak{B}. \tag{6.18}$$

From (6.16) we see that

$$\mathfrak{D}_\theta + \mathfrak{D}_\theta^{-*}\mathfrak{D}^*\mathfrak{D} = (\mathfrak{D}_\theta^{-1})^*. \tag{6.19}$$

The right hand side of (6.19) is an anticausal operator, hence so is the left hand side and so

$$\begin{aligned}
\pi_+(\mathfrak{D}_\theta + \mathfrak{D}_\theta^{-*}\mathfrak{D}^*\mathfrak{D})\pi_- &= 0 \\
\Rightarrow \quad \pi_+\mathfrak{D}_\theta\pi_- &= -\pi_+\mathfrak{D}_\theta^{-*}\mathfrak{D}^*\mathfrak{D}\pi_- \\
&= -\pi_+\mathfrak{D}_\theta^{-*}\mathfrak{D}^*(\pi_+ + \pi_-)\mathfrak{D}\pi_- \\
&= -\pi_+\mathfrak{D}_\theta^{-*}\mathfrak{D}^*\pi_+\mathfrak{D}\pi_-, \quad \text{as } \mathfrak{D}_\theta^{-*}\mathfrak{D}^* \text{ is anticausal,} \\
&= -\pi_+\mathfrak{D}_\theta^{-*}\mathfrak{D}^*\mathfrak{C}\mathfrak{B},
\end{aligned}$$

which is (6.18). Here we have used the well-posedness of Σ, i.e. that $\pi_+\mathfrak{D}\pi_- = \mathfrak{C}\mathfrak{B}$. Note that by our stability assumption (4.6) and the boundedness of $\mathfrak{D}_\theta^{-*}$ it follows from (6.14) that $\mathfrak{C}$ and $\mathfrak{C}_\theta$ are both bounded, and hence so is $\mathfrak{C}_E$. Thus by construction the extended output system $\Sigma_{E_1} = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}_E, \mathfrak{D}_{E_1})$ is a stable L^2 well-posed linear system. The observability Gramian of Σ_{E_1} is given by

$$\begin{aligned}
\mathfrak{C}_E^*\mathfrak{C}_E = \begin{bmatrix} \mathfrak{C}^* & \mathfrak{C}_\theta^* \end{bmatrix} \begin{bmatrix} \mathfrak{C} \\ \mathfrak{C}_\theta \end{bmatrix} &= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}_\theta^*\mathfrak{C}_\theta \\
&= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}\mathfrak{D}_\theta^{-1}\pi_+^2\mathfrak{D}_\theta^{-*}\mathfrak{D}^*\mathfrak{C}, \quad \text{from (6.14),} \\
&= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}(\pi_+ + \pi_-)\mathfrak{D}_\theta^{-1}\pi_+^2\mathfrak{D}_\theta^{-*}(\pi_+ + \pi_-)\mathfrak{D}^*\mathfrak{C} \\
&= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}\pi_+\mathfrak{D}_\theta^{-1}\pi_+^2\mathfrak{D}_\theta^{-*}\pi_+\mathfrak{D}^*\mathfrak{C},
\end{aligned}$$

since $\mathfrak{D}_\theta^{-1}$ is causal and $\mathfrak{D}_\theta^{-*}$ is anticausal. Now an elementary calculation shows that $(\mathfrak{D}_\theta\pi_+)^{-1} = \mathfrak{D}_\theta^{-1}\pi_+$ and thus $(\mathfrak{D}_\theta\pi_+)^{-*} = \pi_+\mathfrak{D}_\theta^{-*}$. Therefore

$$\begin{aligned}
\mathfrak{C}_E^*\mathfrak{C}_E &= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}\pi_+(\mathfrak{D}_\theta\pi_+)^{-1}(\mathfrak{D}_\theta\pi_+)^{-*}\pi_+\mathfrak{D}^*\mathfrak{C} \\
&= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}\pi_+[(\mathfrak{D}_\theta\pi_+)^*(\mathfrak{D}_\theta\pi_+)]^{-1}\pi_+\mathfrak{D}^*\mathfrak{C} \\
&= \mathfrak{C}^*\mathfrak{C} + \mathfrak{C}^*\mathfrak{D}\pi_+[\pi_+\mathfrak{D}_\theta^*\mathfrak{D}_\theta\pi_+]^{-1}\pi_+\mathfrak{D}^*\mathfrak{C} \\
&= P_m, \quad \text{from (6.4) and (6.16).}
\end{aligned}$$

□

Lemma 6.2.5. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ a stable L^2 well-posed realisation. Let $\xi \in H^\infty(\mathbb{C}_0^+; B(\mathscr{Y}))$ denote a spectral factor from Lemma 6.2.3 satisfying* (6.10) *with input-output map $\mathfrak{D}_\xi$. Define*

$$\mathfrak{B}_E := \begin{bmatrix} \mathfrak{B} & \mathfrak{B}_\xi \end{bmatrix} : L^2\left(\mathbb{R}^-; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]\right) \to \mathscr{X}, \tag{6.20}$$

$$\mathfrak{D}_{E_2} := \begin{bmatrix} \mathfrak{D} & \mathfrak{D}_\xi \end{bmatrix} : L^2(\mathbb{R}; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) \to L^2(\mathbb{R}; \mathscr{Y}), \tag{6.21}$$

where

$$\mathfrak{B}_\xi = -\mathfrak{B}\mathfrak{D}^*\mathfrak{D}_\xi^{-*}\pi_- : L^2(\mathbb{R}^-;\mathscr{Y}) \to \mathscr{X}. \tag{6.22}$$

Then $\mathfrak{B}_E$ *is bounded and* $\Sigma_{E_2} := (\mathfrak{A}, \mathfrak{B}_E, \mathfrak{C}, \mathfrak{D}_{E_2})$ *is a stable* L^2 *well-posed linear system on* $(\mathscr{Y}, \mathscr{X}, \left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right])$ *with transfer function*

$$G_{E_2} := \begin{bmatrix} G & \xi \end{bmatrix} \in H^\infty\left(\mathbb{C}_0^+; B\left(\left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right], \mathscr{Y}\right)\right), \tag{6.23}$$

and controllability Gramian Q_m *given by* (6.8)*, which is the optimal cost operator of the dual optimal control problem* (6.5)*.*

Proof. This result essentially follows from Lemma 6.2.4 applied to the dual transfer function and duality. We proceed to give the details. Since the dual transfer function G_d is strictly bounded real, as in Lemma 6.2.3, there exists a spectral factor η, such that $\eta, \eta^{-1} \in H^\infty(\mathbb{C}_0^+; B(\mathscr{Y}))$ and

$$I - [G_d(\mathrm{i}\omega)]^* G_d(\mathrm{i}\omega) = [\eta(\mathrm{i}\omega)]^*\eta(\mathrm{i}\omega), \quad \text{for almost all } \omega \in \mathbb{R}.$$

Let $(\mathfrak{A}^*, {}^d\mathfrak{B}, {}^d\mathfrak{C}, {}^d\mathfrak{D})$ denote the stable L^2 well-posed realisation of G_d from Definition 4.2.1 and let

$$\mathfrak{D}_\eta : L^2(\mathbb{R};\mathscr{Y}) \to L^2(\mathbb{R};\mathscr{Y}),$$

denote the time invariant, causal, bounded, shift-invariant operator corresponding to the transfer function η (again by [81, Theorem 10.3.5]), which is boundedly invertible and satisfies

$$I - ({}^d\mathfrak{D})^* {}^d\mathfrak{D} = \mathfrak{D}_\eta^*\mathfrak{D}_\eta. \tag{6.24}$$

It now follows from the arguments in the proof of Lemma 6.2.4 that $(\mathfrak{A}^*, {}^d\mathfrak{B}, \mathfrak{C}_\eta, \mathfrak{D}_\eta)$ is a stable L^2 well-posed linear system realising η, where

$$\mathfrak{C}_\eta = -\pi_+\mathfrak{D}_\eta^{-*}({}^d\mathfrak{D})^* {}^d\mathfrak{C}. \tag{6.25}$$

Passing to the dual using Definition 4.2.1, we have that $(\mathfrak{A}, \mathfrak{C}_\eta^* R, \mathfrak{C}, {}^d\mathfrak{D}_\eta)$ is a stable L^2 well-posed realisation of $\xi = \eta_d$, where ξ is as in Lemma 6.2.3. We define $\mathfrak{B}_\xi := \mathfrak{C}_\eta^* R$ so that

$$\begin{aligned}\mathfrak{B}_\xi = \mathfrak{C}_\eta^* R &= (-\pi_+\mathfrak{D}_\eta^{-*}({}^d\mathfrak{D})^* {}^d\mathfrak{C})^* R = (-\pi_+\mathfrak{D}_\eta^{-*}(R\mathfrak{D}^*R)^* R\mathfrak{B}^*)^* R \\ &= -\mathfrak{B}\mathfrak{D}^* R\mathfrak{D}_\eta^{-1}\pi_+ R = -\mathfrak{B}\mathfrak{D}^* R\mathfrak{D}_\eta^{-1} R\pi_- = -\mathfrak{B}\mathfrak{D}^*\mathfrak{D}_\xi^{-*}\pi_-,\end{aligned}$$

which is as in (6.22). Here we have used that R is unitary, $\pi_+ R = R\pi_-$ and that

$$\mathfrak{D}_\xi = \mathfrak{D}_{\eta_d} = {}^d\mathfrak{D}_\eta = R\mathfrak{D}_\eta^* R.$$

That $\mathfrak{B}_E$ is bounded follows from the boundedness of the operators in (6.22). Therefore by construction the extended input system $\Sigma_{E_2} = (\mathfrak{A}, \mathfrak{B}_E, \mathfrak{C}, \mathfrak{D}_{E_2})$ is L^2 well-posed with input map $\mathfrak{B}_E$ and input-output map $\mathfrak{D}_{E_2}$ defined by (6.20) and (6.21) respectively. By Lemma 6.2.4, the observability Gramian of

$$\left(\mathfrak{A}^*, {}^d\mathfrak{B}, \begin{bmatrix} {}^d\mathfrak{C} \\ \mathfrak{C}_\eta \end{bmatrix}, \begin{bmatrix} {}^d\mathfrak{D} \\ \mathfrak{D}_\eta \end{bmatrix}\right),$$

is the optimal cost operator Q_m of the dual optimal control problem, which by duality, is the controllability Gramian of the dual system (which by construction is) Σ_{E_2}. $\square$

Remark 6.2.6. For a fixed strictly bounded real transfer function G and stable L^2 well-posed realisation Σ of G there are many extended output systems Σ_{E_1} and many extended input systems Σ_{E_2} owing to the non-uniqueness of the spectral factors θ and ξ from Lemma 6.2.3. However, given any Σ_{E_1}, every other extended output system is determined by Σ_{E_1} and a unitary operator $U \in B(\mathscr{U})$. As such we say that from G and Σ we obtain a family of extended output systems, parameterised by U. Similarly for Σ_{E_2}, now parameterised by unitary $V \in B(\mathscr{Y})$.

6.3 Bounded real balanced truncation

Given a strictly bounded real transfer function G and stable L^2 well-posed realisation of G we now seek to combine an extended output system Σ_{E_1} and an extended input system Σ_{E_2} into one (jointly) extended system with transfer function of the form

$$G_E = \begin{bmatrix} G & \xi \\ \theta & \chi \end{bmatrix},$$

where χ is yet to be determined. We overcome the difficulty of defining χ by making use of an (extended) output map $\mathfrak{C}_E$ and (extended) input map $\mathfrak{B}_E$ from Lemmas 6.2.4 and 6.2.5 above respectively to define a new operator. This operator will turn out to be the Hankel operator corresponding to G_E, which from Definition 5.1.10 determines G_E and hence χ uniquely up to an additive constant. From our putative definition of G_E and Remark 6.2.6 we see that for a given G we will obtain not one, but a family of extended systems, parameterised by two unitary operators. Compare this construction with that in the finite-dimensional case, described in Remark 2.2.4 and Section 2.2.3.

Lemma 6.3.1. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with stable L^2 well-posed realisation Σ. Let θ, ξ denote spectral factors as in Lemma 6.2.3 and let $\mathfrak{C}_E$ and $\mathfrak{B}_E$ denote the output map and input map from Lemma*

6.2.4 and 6.2.5 respectively. Define the bounded operator H_E by

$$H_E := \mathfrak{C}_E \mathfrak{B}_E R : L^2 \left(\mathbb{R}^+; \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix} \right) \to L^2 \left(\mathbb{R}^+; \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix} \right), \tag{6.26}$$

where R is the reflection from Definition 4.2.1. Then H_E is a Hankel operator. The operator H_E is independent of the choice of realisation Σ of G and depends on the spectral factors chosen as follows. If $H_E(\theta_0, \xi_0)$ is the Hankel operator for the choice of spectral factors θ_0, ξ_0 and $H_E(\theta, \xi)$ is the Hankel operator for spectral factors θ, ξ related to θ_0, ξ_0 by (6.11)*, then the Hankel operators are related by*

$$H_E(\theta, \xi) = \begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} H_E(\theta_0, \xi_0) \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix}. \tag{6.27}$$

Remark 6.3.2. In equation (6.27), $\left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix} \right]$ is understood as an operator

$$L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix} \right]) \to L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix} \right]),$$

acting by (pointwise) multiplication. The same is true for $\left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix} \right]$, only now acting on $L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix} \right])$.

Proof of Lemma 6.3.1: For notational convenience let $\mathscr{Z}_1 := \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix} \right]$ and $\mathscr{Z}_2 := \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix} \right]$. Let $v \in L^2(\mathbb{R}^+; \mathscr{Z}_1)$ and $t \geq 0$. It follows from the definition of well-posed linear systems that

$$\mathfrak{B}_E R (\tau_1^t)^* v = \mathfrak{A}^t \mathfrak{B}_E R v. \tag{6.28}$$

Also, for $x \in \mathscr{X}$

$$\mathfrak{C}_E \mathfrak{A}^t x = \tau_2^t \mathfrak{C}_E x. \tag{6.29}$$

Combining (6.28) and (6.29) gives

$$H_E (\tau_1^t)^* = \mathfrak{C}_E \mathfrak{B}_E R (\tau_1^t)^* = \mathfrak{C}_E \mathfrak{A}^t \mathfrak{B}_E R = \tau_2^t \mathfrak{C}_E \mathfrak{B}_E R = \tau_2^t H_E,$$

demonstrating that H_E is Hankel. A calculation shows that

$$H_E = \mathfrak{C}_E \mathfrak{B}_E R = \begin{bmatrix} \mathfrak{C} \\ \mathfrak{C}_\theta \end{bmatrix} \begin{bmatrix} \mathfrak{B} & \mathfrak{B}_\xi \end{bmatrix} R = \begin{bmatrix} \mathfrak{C}\mathfrak{B} & \mathfrak{C}\mathfrak{B}_\xi \\ \mathfrak{C}_\theta \mathfrak{B} & \mathfrak{C}_\theta \mathfrak{B}_\xi \end{bmatrix} R,$$

and using the formulae (6.12) for $\mathfrak{C}_\theta$ and (6.22) for $\mathfrak{B}_\xi$ gives that this equals

$$\begin{bmatrix} \pi_+ \mathfrak{D} \pi_- & \pi_+ \mathfrak{D}_\xi \pi_- \\ \pi_+ \mathfrak{D}_\theta \pi_- & \pi_+ \mathfrak{D}_\theta^{-*} \mathfrak{D}^* \pi_+ \mathfrak{D} \pi_- \mathfrak{D}^* \mathfrak{D}_\xi^{-*} \pi_- \end{bmatrix} R. \tag{6.30}$$

By inspection of (6.30), for given spectral factors θ and ξ, H_E depends only on the terms $\mathfrak{D}, \mathfrak{D}_\theta$, $\mathfrak{D}_\xi$ and their adjoints and inverses where applicable. We recall that an input-output map is completely determined by its transfer function (and vice versa). Therefore, (6.30) depends only on G and the spectral factors θ and ξ. By their construction in Lemma 6.2.3 the spectral factors are certainly independent of the stable L^2 well-posed realisation of G and hence so is H_E.

Equation (6.27) follows from (6.30) and the (easily established) relations

$$\mathfrak{D}_\theta = \mathfrak{D}_{U\theta_0} = U\mathfrak{D}_{\theta_0} \quad \text{and} \quad \mathfrak{D}_\xi = \mathfrak{D}_{\xi_0 V} = \mathfrak{D}_{\xi_0} V.$$

Again U and V are here understood as operators acting on $L^2(\mathbb{R}^+;\mathscr{U})$ and $L^2(\mathbb{R}^+;\mathscr{Y})$ by pointwise multiplication, and certainly commute with π_+, π_- and R. □

Definition 6.3.3. Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U},\mathscr{Y}))$ denote a strictly bounded real transfer function and for a choice of spectral factors θ_0, ξ_0 as in Lemma 6.2.3 let H_E^0 denote the corresponding Hankel operator from Lemma 6.3.1. The set of Hankel operators given by

$$\left\{ \begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} H_E^0 \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} \; : \; U \in B(\mathscr{U}) \text{ unitary, } V \in B(\mathscr{Y}) \text{ unitary.} \right\}, \tag{6.31}$$

is called the family of extended Hankel operators of G.

Remark 6.3.4. It follows from the above definition and the relationships (6.11) and (6.27) that there is a one-to-one correspondence between pairs of spectral factors of G and members of the family of extended Hankel operators of G.

Lemma 6.3.5. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U},\mathscr{Y}))$ denote a strictly bounded real transfer function. Then any two members of the family of extended Hankel operators of G from Definition 6.3.3 have the same singular values. In particular, if one member of this family is nuclear, then all are.*

Proof. Let H_E^0 and H_E denote two members of the family of extended Hankel operators of G which by definition are related by (6.27) for some unitary operators $U \in B(\mathscr{U})$ and $V \in B(\mathscr{Y})$. For notational convenience set $\tilde{U} := \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]$ and $\tilde{V} := \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right]$, so that (6.27) becomes

$$H_E = \tilde{U} H_E^0 \tilde{V}.$$

The operators $\tilde{U}$ and $\tilde{V}$ are unitary and from this an easy calculation shows that for bounded $T : L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]) \to L^2(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right])$

$$\|H_E^0 - T\| = \|\tilde{U} H_E^0 \tilde{V} - \tilde{U} T \tilde{V}\| = \|H_E - \tilde{U} T \tilde{V}\|.$$

It is also easy to see that for any $n \in \mathbb{N}$ the map

$$T \mapsto \tilde{U}T\tilde{V},$$

is a bijection of rank n operators to rank n operators. Therefore for $n \in \mathbb{N}$

$$\begin{aligned} s_n(H_E^0) &= \inf\left\{\|H_E^0 - T\| \,:\, \operatorname{rank} T < n\right\} \\ &= \inf\left\{\|H_E - \tilde{U}T\tilde{V}\| \,:\, \operatorname{rank} T < n\right\} = s_n(H_E). \end{aligned}$$

By counting with multiplicities it follows that $\sigma_k(H_E^0) = \sigma_k(H_E)$ for every $k \in \mathbb{N}$, which completes the proof. ☐

Definition 6.3.6. Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function. We say that G has a nuclear family of extended Hankel operators if some member of the family of extended Hankel operators of G from Definition 6.3.3 is nuclear.

We are now able to construct our desired extended transfer function. As expected, for a fixed original strictly bounded real transfer function we obtain a family of (extended) transfer functions, which we describe in Definition 6.3.8.

Lemma 6.3.7. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function and assume that G has a nuclear family of extended Hankel operators. Let H_E denote a member of this family corresponding to the spectral factors θ, ξ. Then there exists $\chi \in H^\infty(\mathbb{C}_0^+; B(\mathscr{Y}, \mathscr{U}))$ such that*

$$G_E = \begin{bmatrix} G & \xi \\ \theta & \chi \end{bmatrix} \in H^\infty(\mathbb{C}_0^+; B\left(\left[\begin{smallmatrix}\mathscr{U}\\ \mathscr{Y}\end{smallmatrix}\right], \left[\begin{smallmatrix}\mathscr{Y}\\ \mathscr{U}\end{smallmatrix}\right]\right)), \tag{6.32}$$

is regular and is a transfer function of H_E. The feedthrough of χ can without loss of generality be taken equal to zero. Therefore we let

$$D_E = \begin{bmatrix} D & D_\xi \\ D_\theta & 0 \end{bmatrix} : \begin{bmatrix} \mathscr{U} \\ \mathscr{Y} \end{bmatrix} \to \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix}, \tag{6.33}$$

denote the bounded operator such that

$$\lim_{\substack{s \to +\infty \\ s \in \mathbb{R}^+}} G_E(s) = D_E.$$

The components D, D_θ and D_ξ of D_E are the feedthroughs of G, θ and ξ respectively. By always fixing the feedthrough of χ as zero, the Hankel operator H_E and transfer function G_E determine one another uniquely.

Proof. The existence of χ and the regularity of G_E (and hence G and the spectral factors θ and ξ) follows from Corollary 5.1.14. By that result the Hankel operator H_E

determines G_E uniquely up to an additive constant, which we have fixed by demanding that χ has feedthrough zero. $\square$

Definition 6.3.8. Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with nuclear family of extended Hankel operators. By Lemma 6.3.7, each member of this family has a unique transfer function G_E given by (6.32). We call the set of transfer functions G_E the family of extended transfer functions of G.

From its construction, we see that the original transfer function G and the spectral factors θ and ξ are components of the extended transfer function G_E. The next lemma describes how we can obtain L^p well-posed realisations of G and θ from L^p well-posed realisations of G_E, which we shall need later in this work.

Lemma 6.3.9. *Given strictly bounded real $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ with nuclear family of extended Hankel operators, let G_E denote a member of the family of extended transfer functions of G. If $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ is an L^p well-posed realisation on $\left(\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right], \mathscr{X}, \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]\right)$ of G_E with $1 \leq p < \infty$, then*

$$(\mathfrak{A}, \mathfrak{B}|_{\mathscr{U}}, P_{\mathscr{Y}}\mathfrak{C}, P_{\mathscr{Y}}\mathfrak{D}|_{\mathscr{U}}), \quad (\mathfrak{A}, \mathfrak{B}|_{\mathscr{U}}, P_{\mathscr{U}}\mathfrak{C}, P_{\mathscr{U}}\mathfrak{D}|_{\mathscr{U}}), \tag{6.34}$$

are L^p well-posed realisations of G and θ respectively. Here $P_{\mathscr{U}}$ denotes the orthogonal projection of $\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$ onto $\mathscr{U}$ and $P_{\mathscr{Y}}$ denotes the orthogonal projection of $\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$ onto $\mathscr{Y}$. If A, B, C and D denote the generators of $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$, then $A, B|_{\mathscr{U}}, P_{\mathscr{Y}}C$ and $P_{\mathscr{Y}}D|_{\mathscr{U}}$ and $A, B|_{\mathscr{U}}, P_{\mathscr{U}}C$ and $P_{\mathscr{U}}D|_{\mathscr{U}}$ are the generators of the above realisations of G and θ respectively. Furthermore, if $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ is a stable L^2 well-posed realisation of G_E then the realisations in (6.34) are stable L^2 well-posed realisations of G and θ respectively.

Proof. It is routine to verify that the two systems in (6.34) satisfy the conditions of [81, Definition 2.2.1], and hence are L^p well-posed. Since the generators are unique, a short calculation demonstrates that the formulae given are indeed the generators. The final claim is immediate from the definition of a stable L^2 well-posed realisation since restriction and projection are bounded operations. $\square$

We are now in position to define the bounded real singular values and bounded real balanced truncation of a strictly bounded real $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$. The latter, as outlined in Section 6.1, is based on the Lyapunov balanced truncation of a member of the family of extended transfer functions of Definition 6.3.8. As Lemma 6.3.14 demonstrates, however, every member of this family gives rise to the same reduced order transfer function obtained by bounded real balanced truncation.

Definition 6.3.10. Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function. We define the bounded real singular values of G as the singular values (using the convention of Remark 4.3.4) of some member of the family of extended Hankel operators of G from Definition 6.3.3.

Remark 6.3.11.
1. By Lemma 6.3.5 all members of the family of extended Hankel operators of G have the same singular values, so the bounded real singular values depend only on G.
2. By definition, the bounded real singular values are the Lyapunov singular values of any (and hence every) member of the family of extended transfer functions of G.
3. Our next result shows that the above definition is consistent with Definition 2.2.5 of bounded real singular values in the finite-dimensional case. Recall that there the bounded real singular values were defined as the non-negative square roots of the eigenvalues of the product of the bounded real optimal cost operators. An analogous approach in the infinite-dimensional case is trickier because although the product of the optimal cost operators is bounded, it is not *a priori* clear why it should have (non-negative, real) eigenvalues. However, we prove that when the bounded real singular values are summable (which is always true in the finite-dimensional case) then the definitions coincide.

Lemma 6.3.12. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function and let P_m and Q_m denote the optimal cost operators of the optimal control problems* (6.1) *and* (6.5) *corresponding to a given stable L^2 well-posed realisation of G. Then the bounded real singular values are summable if and only if $P_m Q_m$ is compact and the square roots of its eigenvalues are summable. If these conditions hold then apart from possibly zero the bounded real singular values are precisely the square roots of the eigenvalues of $P_m Q_m$ (which therefore depend only on G).*

Proof. Choose a stable L^2 well-posed realisation of G so that P_m and Q_m are given by (6.4) and (6.8) respectively. For some choice of spectral factors as in Lemma 6.2.3, let $\mathfrak{C}_E$ and $\mathfrak{B}_E$ denote the extended output operator and extended input operator from Lemmas 6.2.4 and 6.2.5 respectively. By those results it follows that $P_m = \mathfrak{C}_E^* \mathfrak{C}_E$ and $Q_m = \mathfrak{B}_E \mathfrak{B}_E^*$. Let H_E denote the Hankel operator given by (6.26), which is a member of the family of extended Hankel operators of G. By Definition 6.3.10 the bounded real singular values are the singular values of H_E.

We recall that if H_E is compact then its singular values are precisely the (non-negative) square roots of the eigenvalues of $H_E^* H_E$. We prove that H_E is compact if and only if $P_m Q_m$ is. Assume firstly that H_E is compact, and hence so is

$$P_m Q_m = \mathfrak{C}_E^* \mathfrak{C}_E \mathfrak{B}_E \mathfrak{B}_E^* = \mathfrak{C}_E^* H_E R \mathfrak{B}_E^*,$$

as $\mathfrak{C}_E^*, \mathfrak{B}_E^*$ and R are bounded. Conversely, assume that $P_m Q_m$ is compact. The operator

$$R\mathfrak{B}_E^* P_m Q_m \mathfrak{C}_E^* = R\mathfrak{B}_E^* \mathfrak{C}_E^* \mathfrak{C}_E \mathfrak{B}_E \mathfrak{B}_E^* \mathfrak{C}_E^* = H_E^* H_E H_E^*,$$

is then compact as R, $\mathfrak{B}_E$ and $\mathfrak{C}_E$ are bounded and thus

$$H_E^* H_E H_E^* H_E = (H_E^* H_E)^* H_E^* H_E, \quad \text{is compact.} \tag{6.35}$$

Recall (from, for example, Swartz [83, p. 468]) that the bounded linear operator $T : \mathscr{H}_1 \to \mathscr{H}_2$ ($\mathscr{H}_i$ Hilbert spaces) is compact if and only if T^*T is compact. Invoking this result twice we conclude from (6.35) that H_E is compact. A similar argument shows that $Q_m P_m$ is compact if and only if H_E^* is. Since H_E is compact if and only if H_E^* is (see, for example, Kantorovich & Akilov [42, Theorem IX. 3.3]), we deduce that $P_m Q_m$ is compact if and only if $Q_m P_m$ is.

We now prove that when H_E is compact (equivalently when $P_m Q_m$ or $Q_m P_m$ is compact) that $H_E^* H_E$ and $Q_m P_m$ have the same non-zero eigenvalues with the same multiplicities. If $\mu \neq 0, v \neq 0$ are an eigenvalue, eigenvector pair for $H_E^* H_E$, i.e.

$$H_E^* H_E v = R\mathfrak{B}_E^* \mathfrak{C}_E^* \mathfrak{C}_E \mathfrak{B}_E R v = \mu v, \quad \text{then} \quad Q_m P_m \mathfrak{B}_E R v = \mu \mathfrak{B}_E R v.$$

Since $\mathfrak{B}_E R v \neq 0$, (else $H_E^* H_E v = 0$) it follows that $\mu \in \sigma_p(Q_m P_m)$. Conversely, a similar calculation shows that if $\nu \neq 0$ is an eigenvalue of $Q_m P_m$ with eigenvector w then $R\mathfrak{B}_E^* w \neq 0$ is an eigenvector of $H_E^* H_E$ corresponding to ν and hence $\nu \in \sigma_p(H_E^* H_E)$.

We now prove that the geometric multiplicities of $\mu \neq 0$ as an eigenvalue of $H_E^* H_E$ and $Q_m P_m$ are the same. Suppose that the multiplicities with respect to $H_E^* H_E$ and $Q_m P_m$ are p and q respectively. Choose $(v_i)_{i=1}^p$ linearly independent eigenvectors of $H_E^* H_E$ corresponding to μ. We claim that $(\mathfrak{B}_E R v_i)_{i=1}^p$ (which are eigenvectors of $Q_m P_m$) are linearly independent. If there exist constants $c_i \in \mathbb{C}$ such that

$$\sum_{i=1}^p c_i \mathfrak{B}_E R v_i = 0,$$

then

$$\mu \left(\sum_{i=1}^p c_i v_i \right) = H_E^* H_E \left(\sum_{i=1}^p c_i v_i \right) = H_E^* \mathfrak{C}_E \left(\sum_{i=1}^p c_i \mathfrak{B}_E R v_i \right) = 0.$$

As $\mu \neq 0$ and the v_i are linearly independent, it follows that $c_i = 0$ for each i, which proves the claim and therefore $q \geq p$. Repeating the above argument starting with q linearly independent eigenvectors $(w_i)_{i=1}^q$ of $Q_m P_m$ and proving that $(R\mathfrak{B}_E^* w_i)_{i=1}^q$ are linearly independent eigenvectors of $H_E^* H_E$, we infer that $p \geq q$ and hence $p = q$.

Putting all the above together, if H_E is nuclear, then it is compact and thus so is

$Q_m P_m$. Moreover, the non-zero bounded real singular values are the square roots of the non-zero eigenvalues of $Q_m P_m$, with the same multiplicities, which are therefore summable. Note that $Q_m P_m$ and $P_m Q_m$ have the same eigenvalues (again, apart from possibly zero), with the same multiplicities.

Conversely, assume that $P_m Q_m$ is compact with the square roots of its eigenvalues forming a summable sequence. Then H_E is compact and $H_E^* H_E$ has summable eigenvalues which are, apart from possibly zero, the eigenvalues of $P_m Q_m$. Hence H_E is nuclear and the bounded real singular values are summable.

Finally, from Lemma 6.3.1 it follows that H_E is independent of the stable L^2 well-posed realisation of G chosen, hence so are its singular values and thus when the bounded real singular values are summable, we see that the non-zero eigenvalues of $P_m Q_m$ also depend only on G. $\square$

Definition 6.3.13. Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with summable bounded real singular values and let G_E denote a member of the family of extended transfer functions of G from Definition 6.3.8. Let (A_E, B_E, C_E, D_E) denote the generators from Lemma 5.3.1 of the exactly observable shift realisation ${}^{sr}\Sigma_E^1$ of G_E and for $n \in \mathbb{N}$, let $\mathscr{X}_n$ and $\mathcal{P}_n$ denote the space and projection from Lemma 5.3.4 respectively (defined in terms of the Schimdt vectors of the member of the family of extended Hankel operators corresponding to G_E). Define the operators

$$\begin{aligned} &(A_E)_n := \mathcal{P}_n A|_{\mathscr{X}_n} : \mathscr{X}_n \to \mathscr{X}_n, \quad (B_E)_n^{\mathscr{U}} := \mathcal{P}_n B_E|_{\mathscr{U}} : \mathscr{U} \to \mathscr{X}_n, \\ &(C_E)_n^{\mathscr{Y}} := P_{\mathscr{Y}} C_E|_{\mathscr{X}_n} : \mathscr{X}_n \to \mathscr{Y}. \end{aligned} \tag{6.36}$$

Here $P_{\mathscr{Y}}$ is the orthogonal projection of $\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$ onto $\mathscr{Y}$. We call the finite-dimensional system on $(\mathscr{Y}, \mathscr{X}_n, \mathscr{U})$ generated by $\left[\begin{smallmatrix} (A_E)_n & (B_E)_n^{\mathscr{U}} \\ (C_E)_n^{\mathscr{Y}} & D \end{smallmatrix}\right]$ the reduced order system obtained by bounded real balanced truncation (determined by G_E), where $D = P_{\mathscr{Y}} D_E|_{\mathscr{U}}$ is the feedthrough of G. The function G_n defined by

$$G_n(s) := (C_E)_n^{\mathscr{Y}} (sI - (A_E)_n)^{-1} (B_E)_n^{\mathscr{U}} + D, \tag{6.37}$$

is called the reduced order transfer function obtained from G by bounded real balanced truncation.

Note that the bounded real balanced truncation depends on the choice of extended transfer function G_E. The next lemma shows that different choices of G_E give rise to bounded real balanced truncations of G that are unitarily equivalent to one another. In particular, they all give rise to the same reduced order transfer function in (6.37).

Lemma 6.3.14. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with summable bounded real singular values. Then the bounded real bal-*

anced truncation is unique up to a unitary transformation, determined by the choice of extended transfer function G_E. Every bounded real balanced truncation gives rise to the same reduced order transfer function obtained by bounded real balanced truncation, which is therefore independent of the choice of G_E.

Proof. For the choice of spectral factors θ_0 and ξ_0, let G_E^0 denote the resulting member of the family of extended transfer functions of G. If the spectral factors θ and ξ are related to θ_0 and ξ_0 by (6.11) and G_E is the corresponding extended transfer function, then G_E^0 and G_E are related by

$$G_E = \begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} G_E^0 \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix}, \tag{6.38}$$

where $U \in B(\mathscr{U})$, $V \in B(\mathscr{Y})$ are unitary. The relation (6.38) readily follows from the version for the corresponding extended Hankel operators (6.27) and our definition of the feedthrough D_E of G_E in (6.33).

Let $(A_E^0, B_E^0, C_E^0, D_E^0)$ denote the generators of the exactly observable shift realisation on L^1 of G_E^0. It is readily seen that

$$(A_E, B_E, C_E, D_E) := (A_E^0, \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] B_E^0 \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right], C_E^0, \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] D_E^0 \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right]),$$

generate the exactly observable shift realisation on L^1 of G_E. A simple calculation shows that if $(v_{i,k}^0, w_{i,k}^0)$ are Schmidt pairs of H_E^0 then

$$(v_{i,k}, w_{i,k}) = \left(\left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right]^{-1} v_{i,k}^0, \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] w_{i,k}^0 \right), \tag{6.39}$$

are Schmidt pairs for H_E. Therefore for $n \in \mathbb{N}$

$$\mathscr{X}_n := \langle w_{i,k} \mid 1 \le i \le n,\ 1 \le k \le p_i \rangle = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] \mathscr{X}_n^0,$$

and in fact $\left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] : \mathscr{X}_n^0 \to \mathscr{X}_n$ is an isomorphism. Furthermore if $\mathcal{P}_n^0$ denotes the projection of $L^1(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right])$ onto $\mathscr{X}_n^0$, defined analogously to $\mathcal{P}_n$ in (5.73) then

$$\mathcal{P}_n \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] \mathcal{P}_n^0. \tag{6.40}$$

With these observations we are able to describe how the bounded real balanced truncations of G_E^0 and G_E are related. By definition of the projections $\mathcal{P}_n$, $\mathcal{P}_n^0$ and the operators A_E and A_E^0 we see that for $1 \le i \le n$ and $1 \le k \le p_i$

$$(A_E)_n w_{i,k} = \mathcal{P}_n A_E|_{\mathscr{X}_n} w_{i,k} = \sum_{j=1}^{n} \sum_{r=1}^{p_j} \langle w_{j,r}, \dot{w}_{i,k} \rangle_{L^2} w_{j,r}, \tag{6.41}$$

and

$$\begin{aligned}\left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] (A_E^0)_n \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} w_{i,k} &= \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} \mathcal{P}_n A_E^0|_{\mathscr{X}_n^0} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} w_{i,k} \\ &= \sum_{j=1}^{n} \sum_{r=1}^{p_j} \langle w_{j,r}^0, \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} \dot{w}_{i,k} \rangle_{L^2} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] w_{j,r}^0. \end{aligned} \tag{6.42}$$

The Schmidt pair relations (6.39) and the fact that $\left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]$ is unitary imply that the expressions in (6.41) and (6.42) are equal. Since this equality holds on a basis for $\mathscr{X}_n$ we infer that

$$(A_E)_n = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] (A_E^0)_n \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1}. \tag{6.43}$$

Similarly, using the projection relation (6.40) yields

$$\mathcal{P}_n B_E = \mathcal{P}_n \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] B_E^0 \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right] = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] \mathcal{P}_n^0 B_E^0 \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right],$$

which implies that

$$(B_E)_n^{\mathscr{U}} = \mathcal{P}_n B_E|_{\mathscr{U}} = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] \mathcal{P}_n^0 B_E^0|_{\mathscr{U}} = \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] (B_E^0)_n^{\mathscr{U}}. \tag{6.44}$$

As with A_E^0 and A_E, the operators C_E^0 and C_E are the same and we see that

$$\begin{aligned}(C_E)_n^{\mathscr{Y}} &= P_{\mathscr{Y}} C_E^0|_{\mathscr{X}_n} = P_{\mathscr{Y}} C_E^0|_{\mathscr{X}_n} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} = P_{\mathscr{Y}} C_E^0|_{\mathscr{X}_n^0} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1} \\ &= (C_E^0)_n^{\mathscr{Y}} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]^{-1}. \end{aligned} \tag{6.45}$$

Finally,

$$P_{\mathscr{Y}} D_E^{U,V}|_{\mathscr{U}} = P_{\mathscr{Y}} \left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right] D_E \left[\begin{smallmatrix} I & 0 \\ 0 & V \end{smallmatrix}\right]|_{\mathscr{U}} = D = P_{\mathscr{Y}} D_E|_{\mathscr{U}}. \tag{6.46}$$

We see from (6.43)-(6.46) that the bounded real balanced truncations of $G_E^{U,V}$ and G_E are similar, with unitary similarity transformation $\left[\begin{smallmatrix} I & 0 \\ 0 & U \end{smallmatrix}\right]$. In particular, they both give rise to the same transfer function G_n. □

Our main result of this chapter is the following theorem which describes properties of the reduced order transfer function obtained by bounded real balanced truncation and contains an H^∞ error bound.

Theorem 6.3.15. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function with summable bounded real singular values and where $\mathscr{U}$ and $\mathscr{Y}$ are finite dimensional. Then for each integer n, the reduced order transfer function G_n obtained by bounded real balanced truncation from Definition 6.3.13 is bounded real, rational,* *belongs to* $H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ *and satisfies the bound*

$$\|G - G_n\|_{H^\infty} \leq 2 \sum_{k \geq n+1} \sigma_k, \tag{6.47}$$

where σ_k *are the bounded real singular values.*

In order to prove Theorem 6.3.15 we require a lemma which describes some of the properties of reduced order systems obtained by bounded real balanced truncation. It is the key ingredient in the proof of Theorem 6.3.15, which begins on p. 155.

Lemma 6.3.16. *Let* $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ *denote a strictly bounded real transfer function with summable bounded real singular values and for* $n \in \mathbb{N}$ *let* G_n *denote the reduced order transfer function obtained by bounded real balanced truncation. Then* G_n *is rational, bounded real and for every choice of extended transfer function the resulting bounded real balanced truncation from Definition 6.3.13 is a stable realisation of* G_n*.*

Proof. By Lemma 6.3.14 the bounded real balanced truncations are all unitarily equivalent to one another. Since the stability of A of the realisation $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is invariant under unitary transformation, it does not matter which bounded real balanced truncation we pick. Therefore, for this proof we pick a member G_E of the family of extended transfer functions arbitrarily and for notational convenience we denote the bounded real balanced truncation (depending on G_E) by $\left[\begin{smallmatrix} A_n & B_n^{\mathscr{U}} \\ C_n^{\mathscr{Y}} & D \end{smallmatrix}\right]$.

That G_n is rational is clear, as $\left[\begin{smallmatrix} A_n & B_n^{\mathscr{U}} \\ C_n^{\mathscr{Y}} & D \end{smallmatrix}\right]$ is a realisation on a finite-dimensional state-space. The stability of A_n follows from Proposition 5.3.11 since by construction the Lyapunov balanced truncation of G_E and the bounded real balanced truncation of G have the same main operator A_n.

It remains to see that G_n is bounded real. For this we use the Bounded Real Lemma. We seek a solution of (P, K, W), with $P : \mathscr{X}_n \to \mathscr{X}_n$ self-adjoint and positive, of the bounded real Lur'e equations (2.10) (subject to the realisation $\left[\begin{smallmatrix} A_n & B_n^{\mathscr{U}} \\ C_n^{\mathscr{Y}} & D \end{smallmatrix}\right]$). Noting that

$$C_n = C|_{\mathscr{X}_n} = \begin{bmatrix} P_{\mathscr{Y}} C|_{\mathscr{X}_n} \\ P_{\mathscr{U}} C|_{\mathscr{X}_n} \end{bmatrix} = \begin{bmatrix} C_n^{\mathscr{Y}} \\ C_n^{\mathscr{U}} \end{bmatrix} : \mathscr{X}_n \to \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix}, \tag{6.48}$$

we claim that

$$P = I : \mathscr{X}_n \to \mathscr{X}_n, \qquad K = C_n^{\mathscr{U}} : \mathscr{X}_n \to \mathscr{U}, \qquad W = D_\theta : \mathscr{U} \to \mathscr{U},$$

solve (2.10). The Lyapunov equation (5.110) (considered as an equation on $\mathscr{X}_n$ equipped with the L^2 inner product) from Lemma 5.3.16 can be rewritten using (6.48) as

$$A_n^* + A_n + (C_n^{\mathscr{Y}})^* C_n^{\mathscr{Y}} = -(C_n^{\mathscr{U}})^* C_n^{\mathscr{U}}, \tag{6.49}$$

which is (2.10a). We now prove that the second equation (2.10b) of the bounded real Lur'e equations holds, i.e.

$$B_n^{\mathscr{U}} + (C_n^{\mathscr{Y}})^* D = -(C_n^{\mathscr{U}})^* D_\theta. \tag{6.50}$$

Given stable L^2 well-posed realisations of G and θ, with generators (A, B, C, D) and $(A, B, C_\theta, D_\theta)$ respectively, it is proven in [97, Theorem 12.4] that

$$C_\theta = -D_\theta^{-*}(B_\Lambda^* P_m + D^* C), \quad \text{on } D(A), \tag{6.51}$$

where P_m is the optimal cost operator (6.4), and also that

$$P_m : D(A) \to D(B_\Lambda^*). \tag{6.52}$$

The operator B_Λ^* in (6.51) is the Λ-extension of B^*, given by

$$B_\Lambda^* x = \lim_{\substack{\alpha \to +\infty \\ \alpha \in \mathbb{R}^+}} B^* \alpha (\alpha I - A^*)^{-1} x,$$

with domain consisting of the $x \in \mathscr{X}$ such that the above limit exists. We consider equality (6.51) and the condition (6.52) for the generators of the (stable L^2 well-posed) realisations of G and θ obtained using Lemma 6.3.9 from the output-normal realisation ${}^{sr}\Sigma_E^2$ of G_E. The observability Gramian of ${}^{sr}\Sigma_E^2$ is the identity. Therefore, by Lemma 6.2.4, P_m also equals the identity. We let $(\mathcal{A}_E, \mathcal{B}_E, \mathcal{C}_E, D_E)$ denote the generators of ${}^{sr}\Sigma_E^2$, which are given by Lemma 5.3.1. We note from that result that the operator $\mathcal{B}_E$ does not map into L^2, as h_E may not belong to L^2 necessarily. From Lemma 6.3.9 it follows that $(\mathcal{A}_E, \mathcal{B}_E|_{\mathscr{U}}, P_{\mathscr{Y}}\mathcal{C}_E, P_{\mathscr{Y}} D_E|_{\mathscr{U}})$ and $(\mathcal{A}_E, \mathcal{B}_E|_{\mathscr{U}}, P_{\mathscr{U}}\mathcal{C}_E, P_{\mathscr{U}} D_E|_{\mathscr{U}})$, which for notational convenience we denote by $(\mathcal{A}, \mathcal{B}_{\mathscr{U}}, \mathcal{C}^{\mathscr{Y}}, D)$ and $(\mathcal{A}, \mathcal{B}_{\mathscr{U}}, \mathcal{C}_\theta^{\mathscr{U}}, D_\theta)$, are the generators of stable L^2 well-posed realisations of G and θ respectively. Rewriting (6.51) in terms of these above generators yields

$$-(D^* \mathcal{C}^{\mathscr{Y}} + D_\theta^* \mathcal{C}_\theta^{\mathscr{U}}) = (\mathcal{B}_{\mathscr{U}}^*)_\Lambda \cdot I = (\mathcal{B}_{\mathscr{U}}^*)_\Lambda, \quad \text{on } D(\mathcal{A}) = W^{1,2}. \tag{6.53}$$

The condition (6.52) implies that

$$D(\mathcal{A}) = W^{1,2}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]) \subseteq D((\mathcal{B}_{\mathscr{U}}^*)_\Lambda).$$

We prove first of all that

$$(\mathcal{B}_{\mathscr{U}}^*)_\Lambda x = P_{\mathscr{U}}(H_E^* x)(0), \quad \forall\, x \in D(\mathcal{A}). \tag{6.54}$$

To that end, consider for $x \in D(\mathcal{A})$ and $u \in \mathscr{U}$ (which is understood as $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$)

$$\begin{aligned} \langle \mathcal{B}_\Lambda^* x, u \rangle_{\mathscr{U}} &= \langle \lim_{\substack{\alpha \to \infty \\ \alpha \in \mathbb{R}^+}} \mathcal{B}^* \alpha (\alpha I - \mathcal{A}^*)^{-1} x, u \rangle_{\mathscr{U}} \\ &= \lim_{\substack{\alpha \to \infty \\ \alpha \in \mathbb{R}^+}} \langle x, \alpha(\alpha I - \mathcal{A})^{-1} \mathcal{B} u \rangle_{L^2}. \end{aligned} \tag{6.55}$$

By [81, Example 4.2.6 (ii)] for $\alpha \in \mathbb{R}^+$ and $u \in \mathscr{U}$

$$\alpha(\alpha I - \mathcal{A})^{-1} \mathcal{B} u = H_E(\alpha \mathrm{e}_{-\alpha} u), \tag{6.56}$$

where $(\alpha \mathrm{e}_{-\alpha} u)(t) = \alpha \mathrm{e}^{-\alpha t} u$, and so combining (6.55) and (6.56) gives

$$\langle \mathcal{B}_\Lambda^* x, u \rangle_{\mathscr{U}} = \lim_{\substack{\alpha \to \infty \\ \alpha \in \mathbb{R}^+}} \langle x, H_E(\alpha \mathrm{e}_{-\alpha} u) \rangle_{L^2}. \tag{6.57}$$

Now consider

$$\langle x, H_E(\alpha \mathrm{e}_{-\alpha} u) \rangle_{L^2} = \langle H_E^* x, \alpha \mathrm{e}_{-\alpha} u \rangle_{L^2},$$

which we integrate by parts to give

$$\begin{aligned} &= \left[\langle (H_E^* x)(t), \mathrm{e}^{-\alpha t} u \rangle_{\mathscr{U}} \right]_\infty^0 + \left\langle \frac{d}{dt}(H_E^* x), \mathrm{e}_{-\alpha} u \right\rangle_{L^2} \\ &= \langle (H_E^* x)(0), u \rangle_{\mathscr{U}} - \langle h_E^* x(0), \mathrm{e}_{-\alpha} u \rangle_{L^2} \\ &\quad - \langle H_E^* \dot{x}, \mathrm{e}_{-\alpha} u \rangle_{L^2}. \end{aligned} \tag{6.58}$$

We investigate the limiting behaviour as $\alpha \to \infty$ of the second and third term on the right-hand side of (6.58) separately. We have

$$\begin{aligned} \langle h_E^* x(0), \mathrm{e}_{-\alpha} u \rangle_{L^2} &= \int_{\mathbb{R}^+} \left\langle h_E^*(t) x(0), \mathrm{e}^{-\alpha t} u \right\rangle_{\mathscr{U}} \, dt \\ &\leq \|u\|_{\mathscr{U}} \|x(0)\| \int_{\mathbb{R}^+} \|h_E^*(t)\| \mathrm{e}^{-\alpha t} \, dt \to 0, \quad \text{as} \quad \alpha \to \infty, \end{aligned} \tag{6.59}$$

by the Dominated Convergence Theorem, since

$$\mathbb{R}^+ \ni t \mapsto h_E^*(t) \in L^1, \quad \text{and} \quad \lim_{\alpha \to +\infty} \|h_E^*(t)\| \mathrm{e}^{-\alpha t} = 0, \quad a.a.\ t \geq 0.$$

Observe that H_E and H_E^* satisfy the assumption **A** from Section 5.2 and so the bound (5.27) from Lemma 5.2.1 holds. Therefore, by the Cauchy-Schwarz inequality on L^2

$$\begin{aligned} \langle H_E^* \dot{x}, \mathrm{e}_{-\alpha} u \rangle_{L^2} &\leq \|H_E^* \dot{x}\|_2 \cdot \|\mathrm{e}_{-\alpha} u\|_2 \\ &\leq \|h_E^*\|_1 \cdot \|\dot{x}\|_2 \cdot \|\mathrm{e}_{-\alpha} u\|_2 \to 0, \quad \text{as} \quad \alpha \to \infty, \end{aligned} \tag{6.60}$$

where we have used (5.27) and the fact that

$$\|e_{-\alpha}u\|_2^2 = \|u\|_{\mathscr{U}}^2 \int_{\mathbb{R}^+} e^{-2\alpha t}\, dt = \frac{1}{2\alpha}\|u\|_{\mathscr{U}}^2 \to 0, \quad \text{as} \quad \alpha \to \infty.$$

By substituting (6.59) and (6.60) into (6.58) and taking the limit $\alpha \to \infty$, we infer from (6.57) that

$$\langle \mathcal{B}_\Lambda^* x, \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{U}\times\mathscr{Y}} = \langle (H_E^* x)(0), \left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]\rangle_{\mathscr{U}\times\mathscr{Y}},$$

and as $u \in \mathscr{U}$ was arbitrary, we conclude that (6.54) holds.

Next observe that the left-hand side of (6.53) can be written as

$$-P_{\mathscr{U}} \begin{bmatrix} D^* & D_\theta^* \\ D_\xi^* & 0 \end{bmatrix} \begin{bmatrix} \mathcal{C}^{\mathscr{Y}} \\ \mathcal{C}_\theta^{\mathscr{U}} \end{bmatrix} = -P_{\mathscr{U}} D_E^* \mathcal{C}_E, \tag{6.61}$$

where $\mathcal{C}_E$ is given by (5.67). Combining (6.53), (6.54) and (6.61) gives

$$-P_{\mathscr{U}} D_E^* \mathcal{C}_E x = P_{\mathscr{U}}(H_E^* x)(0), \quad \forall\, x \in W^{1,2}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]). \tag{6.62}$$

Equality (6.62) is a relationship the generators of ${}^{sr}\Sigma_E^2$ satisfy. We seek to prove that the corresponding generators of the realisation ${}^{sr}\Sigma_E^1$ satisfy the same relationship. We do this by proving that (6.62) can be extended to an equality on $W^{1,1}$. To that end restrict (6.62) to $W^{1,1} \cap W^{1,2} \subseteq W^{1,2}$, i.e.

$$-P_{\mathscr{U}} D_E^* \mathcal{C}_E y = P_{\mathscr{U}}(H_E^* y)(0), \quad \forall\, y \in W^{1,1} \cap W^{1,2}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]). \tag{6.63}$$

and observe that $W^{1,1} \cap W^{1,2}$ is dense in $W^{1,1}$ (since the set of test functions with compact support on $\mathbb{R}^+$ is contained in $W^{1,1} \cap W^{1,2}$ and is dense in $W^{1,1}$). As both sides of (6.63) are continuous on $W^{1,1}$, we can extend (6.63) by density and continuity to

$$-P_{\mathscr{U}} D_E^* y(0) = P_{\mathscr{U}}(H_E^* y)(0), \quad \forall\, y \in W^{1,1}(\mathbb{R}^+; \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]). \tag{6.64}$$

Since $w_{i,k} \in W^{1,1}$, from equation (6.64) we see that

$$\begin{aligned} &-P_{\mathscr{U}} D_E^* w_{i,k}(0) = P_{\mathscr{U}}(H_E^* w_{i,k})(0), \quad \forall\, 1 \le i \le n, \quad \forall\, 1 \le k \le p_i, \\ \Rightarrow \quad &-P_{\mathscr{U}} D_E^* C_n = (B_n^{\mathscr{U}})^*, \quad \text{on } \mathscr{X}_n. \end{aligned} \tag{6.65}$$

In the above we have used that

$$B_n^{\mathscr{U}} = P_{\mathscr{X}_n}(B_E)|_{\mathscr{U}} : \mathscr{U} \to \mathscr{X}_n, \quad C_n = \begin{bmatrix} C_n^{\mathscr{Y}} \\ C_n^{\mathscr{U}} \end{bmatrix} = \begin{bmatrix} P_{\mathscr{Y}} C_E \\ P_{\mathscr{U}} C_E \end{bmatrix} |_{\mathscr{X}_n} : \mathscr{X}_n \to \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix},$$

where B_E and C_E are generators of the realisation ${}^{sr}\Sigma_E^1$. Note that they act in the same way as the operators in (5.68) an (5.67), but have different domains and codomains.

In (6.65) we have also used Lemma 5.3.3 to infer that

$$(B_n^{\mathscr{U}})^* : \mathscr{X}_n \to \mathscr{U}, \quad (B_n^{\mathscr{U}})^* w_{i,k} = P_{\mathscr{U}}(H_E^* w_{i,k})(0).$$

Inserting $D_E = \begin{bmatrix} D & D_\xi \\ D_\theta & 0 \end{bmatrix}$ into equation (6.65) gives

$$(B_n^{\mathscr{U}})^* + D^* C_n^{\mathscr{Y}} = -D_\theta^* C_n^{\mathscr{U}},$$

as an operator equation from $\mathscr{X}_n$ to $\mathscr{Y}$ (both finite-dimensional), which when adjointed gives (6.50), as required.

It remains to prove that the third equation (2.10c) of the bounded real Lur'e equations holds, i.e.

$$I - D^* D = D_\theta^* D_\theta, \tag{6.66}$$

In Staffans [78, Corollary 7.2] (see also [97, Remark 12.9]) the following formula is given relating the feedthroughs of the original transfer function G and the spectral factor θ (both of which are regular by Lemma 6.3.7):

$$D_\theta^* D_\theta = I - D^* D + \lim_{\substack{s\to\infty \\ s\in\mathbb{R}^+}} B_\Lambda^* P_m (sI - A)^{-1} B. \tag{6.67}$$

Equality (6.67) holds for any stable (Hilbert space state-space) realisation of G and θ. We consider the stable L^2 well-posed realisation of G with generators $(\mathcal{A}, \mathcal{B}_{\mathscr{U}}, \mathcal{C}^{\mathscr{Y}}, D)$ (i.e. the realisation obtained from the output-normal realisation of G_E by Lemma 6.3.9). Proving that equality (6.66) holds is equivalent to proving that

$$\begin{aligned} \lim_{\substack{s\to\infty \\ s\in\mathbb{R}^+}} (\mathcal{B}_{\mathscr{U}}^*)_\Lambda (sI - \mathcal{A})^{-1} \mathcal{B}_{\mathscr{U}} &= \lim_{\substack{s\to\infty \\ s\in\mathbb{R}^+}} \lim_{\substack{\alpha\to\infty \\ \alpha\in\mathbb{R}^+}} \mathcal{B}_{\mathscr{U}}^* \alpha(\alpha I - \mathcal{A}^*)^{-1} (sI - \mathcal{A})^{-1} \mathcal{B}_{\mathscr{U}} \\ &= 0. \end{aligned} \tag{6.68}$$

We consider for $u, v \in \mathscr{U}$ (again really $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} v \\ 0 \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right]$) and $\alpha, s \in \mathbb{R}^+$ sufficiently large

$$\langle \mathcal{B}^* \alpha(\alpha I - \mathcal{A}^*)^{-1} (sI - \mathcal{A})^{-1} \mathcal{B} u, v \rangle_{\mathscr{U}} = \langle (sI - \mathcal{A})^{-1} \mathcal{B} u, \alpha(\alpha I - \mathcal{A})^{-1} \mathcal{B} v \rangle_{L^2}$$

and recall (6.56) which gives that $(sI - \mathcal{A})^{-1}\mathcal{B}u = H_E(\mathrm{e}_{-s}u)$, where $(\mathrm{e}_{-s}u)(t) = \mathrm{e}^{-st}u$. Thus

$$\begin{aligned} |\langle \mathcal{B}^* \alpha(\alpha I - \mathcal{A}^*)^{-1} (sI - \mathcal{A})^{-1} \mathcal{B} u, v \rangle_{\mathscr{U}}| &= |\langle H_E(\mathrm{e}_{-s}u), H_E(\alpha \mathrm{e}_{-\alpha} v) \rangle_{L^2}| \\ &\le \|H_E(\mathrm{e}_{-s}u)\|_\infty \|H_E(\alpha \mathrm{e}_{-\alpha} u)\|_1, \end{aligned} \tag{6.69}$$

by the Hölder inequality. Note that for $\alpha \in \mathbb{R}^+$

$$\alpha \mathrm{e}_{-\alpha} u \in L^1 \quad \text{as} \quad \|\alpha \mathrm{e}_{-\alpha} u\|_1 = \|u\|_{\mathscr{U}},$$

and so

$$\|H_E(\alpha \mathrm{e}_{-\alpha} u)\|_1 \leq \|h_E\|_1 \cdot \|\alpha \mathrm{e}_{-\alpha} u\|_1 = \|h_E\|_1 \cdot \|u\|_{\mathscr{U}}, \tag{6.70}$$

where we have used (5.27). We now prove that

$$\|H_E(\mathrm{e}_{-s} u)\|_\infty \to 0, \quad \text{as} \quad s \to \infty. \tag{6.71}$$

Recall the expression for the kernel h_E from (5.23) of the nuclear Hankel operator H_E

$$h_E(\tau) = \sum_{n \in \mathbb{N}_0} \lambda_n (\mathrm{Re}\ a_n) \mathrm{e}^{a_n \tau}, \quad \tau > 0.$$

Observe that for $t, \tau \geq 0$

$$\|h_E(t+\tau)\mathrm{e}^{-s\tau}\| \leq \sum_{n \in \mathbb{N}_0} \|\lambda_n\| (-\mathrm{Re}\ a_n) \mathrm{e}^{\mathrm{Re}\ a_n \tau} \mathrm{e}^{-s\tau} = f(\tau)\mathrm{e}^{-s\tau},$$

where

$$\tau \mapsto f(\tau) := \sum_{n \in \mathbb{N}_0} \|\lambda_n\| (-\mathrm{Re}\ a_n) \mathrm{e}^{\mathrm{Re}\ a_n \tau} \in L^1.$$

Therefore

$$\|H_E(\mathrm{e}_{-s} u)(t)\| = \int_{\mathbb{R}^+} \|h_E(t+\tau)\mathrm{e}^{-s\tau}\|\, d\tau \leq \int_{\mathbb{R}^+} f(\tau)\mathrm{e}^{-s\tau}\, d\tau \to 0, \quad \text{as} \quad s \to \infty,$$

by the Dominated Convergence Theorem. Since $t \geq 0$ was arbitrary it follows that (6.71) holds as required. Inserting (6.70) and (6.71) into (6.69) and taking the limits $\alpha \to \infty$ then $s \to \infty$, we conclude that

$$\left\langle \lim_{\substack{s \to \infty \\ s \in \mathbb{R}^+}} (\mathcal{B}_{\mathscr{U}}^*)_\Lambda (sI - \mathcal{A})^{-1} \mathcal{B}_{\mathscr{U}} u, v \right\rangle_{\mathscr{U}} = 0.$$

As $u, v \in \mathscr{U}$ were arbitrary we deduce (6.68) holds and hence equality (6.66) is established. Therefore we have proven

$$\begin{aligned} A_n^* + A_n + (C_n^{\mathscr{Y}})^* C_n^{\mathscr{Y}} &= -(C_n^{\mathscr{U}})^* C_n^{\mathscr{U}}, \\ B_n^{\mathscr{U}} + (C_n^{\mathscr{Y}})^* D &= -(C_n^{\mathscr{U}})^* D_\theta, \\ I - D^* D &= D_\theta^* D_\theta, \end{aligned}$$

which states that the identity $I : \mathscr{X}_n \to \mathscr{X}_n$ is a (self-adjoint, positive) solution of the bounded real Lur'e equations and hence G_n is bounded real. $\square$

Remark 6.3.17. When the limit in (6.68) is zero has been considered for different types of systems in the PhD thesis of Mikkola [50]. That the limit is zero in our case is known and is contained in [50, Theorem 9.1.15], but we have given a proof here for convenience.

We now have all the ingredients to prove Theorem 6.3.15.

Proof of Theorem 6.3.15: That the reduced order transfer function obtained by bounded real balanced truncation G_n is rational and bounded real follows from Lemma 6.3.16. It remains to prove the error bound (6.47). By Lemma 6.3.5, every Hankel operator H_E of an extended system Σ_E with transfer function G_E, is nuclear. So Theorem 5.0.2 applied to G_E yields

$$\|G_E - (G_E)_n\|_{H^\infty} \leq 2 \sum_{k=n+1}^{\infty} \sigma_k, \tag{6.72}$$

where $(G_E)_n$ is the Lyapunov balanced truncation of G_E (*not* the bounded real balanced truncation), and σ_k are the Lyapunov singular values of G_E and so are also the bounded real singular values of G, by Definition 6.3.10. By construction of G_E in (6.32) we have that

$$G(s) = P_{\mathscr{Y}} G_E(s)|_{\mathscr{U}}.$$

Moreover, by construction of the bounded real balanced truncation and Lyapunov balanced truncation (see equations (6.37) and (5.77) from Definitions 6.3.13 and 5.3.5 respectively)

$$G_n(s) = P_{\mathscr{Y}} (G_E)_n(s)|_{\mathscr{U}}.$$

Together these yield

$$\|G - G_n\|_{H^\infty} = \|P_{\mathscr{Y}}(G_E - (G_E)_n)|_{\mathscr{U}}\|_{H^\infty} \leq \|G_E - (G_E)_n\|_{H^\infty}. \tag{6.73}$$

Combining (6.72) and (6.73) gives the result. □

6.4 Notes

The aim of this chapter is to extend model reduction by bounded real balanced truncation to infinite-dimensional systems. The main result of this chapter is Theorem 6.3.15, which is new. A version of this chapter, together with material from the next two chapters, has been submitted for publication as [37]. As outlined in Section 6.1 our approach has been to view bounded real balanced truncation as Lyapunov balanced truncation of a certain extended system. The key ingredients in construction of the extended system are the optimal control results of [97] and [77] and the extended (spectral factor) systems of [97]. Lyapunov balanced truncation has been described in Chapter 5. We comment on some possible extensions and future work in Chapter 9.

Chapter 7

Positive real balanced truncation

In this chapter we extend positive real balanced truncation to a class of infinite-dimensional systems. The main result of this chapter is Theorem 7.2.12 which states that for strictly positive real J with summable positive real singular values $(\sigma_k)_{k\in\mathbb{N}}$ for each $n \in \mathbb{N}$ there exists a rational positive real transfer function J_n such that the error bound

$$\hat{\delta}(J, J_n) \leq 2 \sum_{k=n+1}^{N} \sigma_k,$$

holds. The results of this chapter are largely proven by applying the Cayley transform, described below in Section 7.1, to the bounded real balanced truncation of Chapter 6. The motivation for extending positive real balanced truncation is also the same as in bounded real case- because of the (often) very rapid rates of decay of the positive real singular values, resulting in much faster convergence of the positive real balanced truncation compared to other numerical schemes. We give an example highlighting these varying convergence rates in Chapter 8.

7.1 The Cayley transform

As is well-known, bounded real and positive real systems are related by the Cayley transform (also known as the diagonal transform or Möbius transform). Here we collect the material we will need in order to be able to convert bounded real balanced truncation to positive real balanced truncation. We start by recalling some known definitions.

Definition 7.1.1. For $\mathscr{X}$ a Hilbert space define the set

$$D(\mathcal{S}_{\mathscr{X}}) := \{T \in B(\mathscr{X}) \ : \ -1 \in \rho(T)\}.$$

The map $\mathcal{S}_{\mathscr{X}} : D(\mathcal{S}_{\mathscr{X}}) \to D(\mathcal{S}_{\mathscr{X}})$ given by

$$D(\mathcal{S}_{\mathscr{X}}) \ni T \mapsto \mathcal{S}_{\mathscr{X}}(T) := (I - T)(I + T)^{-1} \in B(\mathscr{X}),$$

is the Cayley transform. It is self-inverse.

Remark 7.1.2. For notational convenience we define for $\mathscr{U}$ a Hilbert space

$$\mathcal{S} = \mathcal{S}_{\mathscr{U}}, \quad \check{\mathcal{S}} = \mathcal{S}_{L^2(\mathbb{R}^+;\mathscr{U})}.$$

Definition 7.1.3. For $\mathscr{U}$ a Hilbert space define the set

$$D(\tilde{\mathcal{S}}) := \left\{ G : \mathbb{C}_0^+ \to B(\mathscr{U}) \ : \ -1 \in \rho(G(s)), \ \forall\, s \in \mathbb{C}_0^+ \right\}.$$

We also call the map $\tilde{\mathcal{S}} : D(\tilde{\mathcal{S}}) \to D(\tilde{\mathcal{S}})$ defined by

$$D(\tilde{\mathcal{S}}) \ni G \mapsto \Big(\mathbb{C}_0^+ \ni s \mapsto [\tilde{\mathcal{S}}(G)](s) := \mathcal{S}(G(s)) \Big),$$

the Cayley transform. It is also self-inverse.

Remark 7.1.4. The term Cayley transform (often, with parameter $\alpha \in \mathbb{C}$) is sometimes used in the literature to denote the transform

$$G \mapsto \tilde{G} = \left(\mathbb{D} \ni z \mapsto \tilde{G}(z) = G\left(\frac{1-z}{1+z} \right) \right). \tag{7.1}$$

Since the function $s \mapsto \frac{1-s}{1+s}$ is a bijection between $\mathbb{D}$ and $\mathbb{C}_0^+$, the transformation in (7.1) can be considered as a mapping from a continuous time system (often a system node) to a discrete time system. See for example Staffans [80, Section 7]. We remark that this is not our interpretation here.

Example 7.1.5. We note that functions in $D(\tilde{\mathcal{S}})$ need not belong to $H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, for example the positive real, rational function J

$$\mathbb{C}_0^+ \ni s \mapsto J(s) := s, \tag{7.2}$$

certainly satisfies $J \in D(\tilde{\mathcal{S}})$, but J is unbounded and so $J \notin H^\infty(\mathbb{C}_0^+)$. We infer that $D(\tilde{\mathcal{S}}) \not\subseteq H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$

Example 7.1.6. Not every bounded real function belongs to $D(\tilde{\mathcal{S}})$; the function

$$G(s) \equiv -I,$$

is bounded real, but $G \notin D(\tilde{\mathcal{S}})$.

In order to carefully describe how the Cayley transform maps bounded real functions to positive real functions (and vice versa) we need the following lemma.

Lemma 7.1.7. *Let $\mathscr{U}$ denote a finite-dimensional Hilbert space. If $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ is strictly bounded real then there exist positive constants c_1 and C_1 such that*

$$c_1\|u\|_{\mathscr{U}} \leq \|(I+G(s))u\|_{\mathscr{U}} \leq C_1\|u\|_{\mathscr{U}}, \quad \forall\, s \in \mathbb{C}_0^+, \quad \forall\, u \in \mathscr{U}. \tag{7.3}$$

Therefore $(I+G(s))^{-1}$ exists for all $s \in \mathbb{C}_0^+$ and there exists constants c_2 and C_2 such that

$$c_2\|u\|_{\mathscr{U}} \leq \|(I+G(s))^{-1}u\|_{\mathscr{U}} \leq C_2\|u\|_{\mathscr{U}}, \quad \forall\, s \in \mathbb{C}_0^+, \quad \forall\, u \in \mathscr{U}. \tag{7.4}$$

If $J : \mathbb{C}_0^+ \to B(\mathscr{U})$ is positive real then $(I+J(s))^{-1}$ exists for all $s \in \mathbb{C}_0^+$ and the following bounds hold for all $s \in \mathbb{C}_0^+$ and all $u \in \mathscr{U}$

$$\|u\|_{\mathscr{U}} \leq \|(I+J(s))u\|_{\mathscr{U}}, \tag{7.5}$$

$$\|(I+J(s))^{-1}u\|_{\mathscr{U}} \leq \|u\|_{\mathscr{U}}. \tag{7.6}$$

If additionally $J \in H^\infty(\mathbb{C}_0^+)$ then there exist constants c_3, C_3 such that

$$\|u\|_{\mathscr{U}} \leq \|(I+J(s))u\|_{\mathscr{U}} \leq C_3\|u\|_{\mathscr{U}}, \tag{7.7}$$

$$c_3\|u\|_{\mathscr{U}} \leq \|(I+J(s))^{-1}u\|_{\mathscr{U}} \leq \|u\|_{\mathscr{U}}, \tag{7.8}$$

for all $s \in \mathbb{C}_0^+$ and all $u \in \mathscr{U}$. All of the above inequalities also hold for almost all $s \in \mathrm{i}\mathbb{R}$ and all $u \in \mathscr{U}$.

Proof. An application of the triangle inequality and the reverse triangle inequality using (2.9) for strictly bounded real G readily gives

$$\varepsilon\|u\|_{\mathscr{U}} \leq \|(I+G(s))u\|_{\mathscr{U}} \leq (2-\varepsilon)\|u\|_{\mathscr{U}},$$

which is (7.3). The lower bound in (7.3) implies that $I+G(s)$ is injective, and thus invertible (both uniformly in s) as $\mathscr{U}$ is finite-dimensional. The inequalities in (7.4) now follow from those in (7.3).

Now suppose J is positive real. For $s \in \mathbb{C}_0^+$ and $u \in \mathscr{U}$ we have that

$$\begin{aligned}\|(I+J(s))u\|_{\mathscr{U}}^2 &= \langle (I+J(s))u, (I+J(s))u\rangle_{\mathscr{U}} \\ &= \langle u, u\rangle_{\mathscr{U}} + \underbrace{\langle J(s)u, J(s)u\rangle_{\mathscr{U}}}_{\geq 0} + \underbrace{\langle (J(s)+(J(s))^*)u, u\rangle_{\mathscr{U}}}_{\geq 0} \\ &\geq \|u\|_{\mathscr{U}}^2,\end{aligned}$$

where we have used the positive realness of J. We conclude that (7.5) holds, implying that $(I+J(s))^{-1}$ exists (for the same reasons as in the bounded real case) and is bounded (uniformly in s), which is (7.6).

Now assume additionally that $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, then by the triangle inequality for $s \in \mathbb{C}_0^+$ and $u \in \mathscr{U}$

$$\|(I + J(s))u\|_{\mathscr{U}} \leq \|u\|_{\mathscr{U}} + \|J(s)u\|_{\mathscr{U}} \leq (1 + \|J\|_{H^\infty})\|u\|_{\mathscr{U}},$$

which when combined with (7.5) gives (7.7). Finally, (7.8) follows from (7.7). The versions of the inequalities on the imaginary axis hold by taking limits. $\square$

Lemma 7.1.8. *Given the Cayley transform $\tilde{\mathcal{S}}$ of Definition 7.1.3, and $\mathscr{U}$ a finite-dimensional Hilbert space, let BR, PR, SBR and SPR denote the sets of functions $\mathbb{C}_0^+ \to B(\mathscr{U})$ that are bounded real, positive real, strictly bounded real or strictly positive real respectively. Then*

(i) $BR \not\subseteq D(\tilde{\mathcal{S}})$.

(ii) $PR \subseteq D(\tilde{\mathcal{S}})$.

(iii) $SBR \subseteq D(\tilde{\mathcal{S}})$.

(iv) $\tilde{\mathcal{S}} : SBR \to H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$.

(v) $\tilde{S} : BR \cap D(\tilde{S}) \to PR$ *is a bijection.*

(vi) $\tilde{\mathcal{S}} : SBR \to SPR$.

(vii) $\tilde{\mathcal{S}} : SPR \cap H^\infty \to SBR$ *is a bijection.*

Proof. Claim (i) follows from Example 7.1.5.

Observe that if $(I + G(s))^{-1}$ exists for all $s \in \mathbb{C}_0^+$ and is uniformly bounded in s, then $G \in D(\tilde{\mathcal{S}})$. Claims (ii) and (iii) follow from inequalities (7.6) and (7.4) from Lemma 7.1.7 respectively and the previous remark.

Claim (iv) follows from the fact that $\tilde{\mathcal{S}}$ clearly preserves analyticity and the inequalities

$$\|\tilde{\mathcal{S}}(G)(s)u\|_{\mathscr{U}} = \|(I - G(s))(I + G(s))^{-1}u\| \leq C_1(1 + \|G\|_{H^\infty})\|u\|_{\mathscr{U}},$$

where we have used (7.4), which demonstrate that $\tilde{\mathcal{S}}(G)$ is bounded on the open right-half plane.

The claims $(v) - (vii)$ describe how the Cayley transform maps bounded real functions into positive real and vice versa. These results are known and are similar to the arguments in, for example, in Belevitch [9, p.160, 189].

We need the following characterisations of bounded realness and strictly bounded realness. It is easy to see from the definition of bounded real that a function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ is bounded real if and only if for all $s \in \mathbb{C}_0^+$

$$I - [G(s)]^* G(s) \geq 0. \tag{7.9}$$

Similarly, G is strictly bounded real if and only if there exists an $\varepsilon > 0$ such that for all $s \in \mathbb{C}_0^+$

$$I - [G(s)]^* G(s) \geq \varepsilon I. \tag{7.10}$$

The above inequalities (7.9) and (7.10) also hold for almost all $s \in \mathrm{i}\mathbb{R}$ when G is bounded real or strictly bounded real, respectively.

(v): Suppose $G \in BR \cap D(\tilde{\mathcal{S}})$ and let $J := \tilde{\mathcal{S}}(G)$. We calculate

$$\begin{aligned} J + J^* &= (I+G)^{-*}(I-G)^* + (I-G)(I+G)^{-1} \\ &= (I+G)^{-*}\left[(I-G)^*(I+G) + (I+G)^*(I-G)\right](I+G)^{-1} \\ &= 2(I+G)^{-*}[I - G^*G](I+G)^{-1}. \end{aligned}$$

Therefore for $s \in \mathbb{C}_0^+$ and $u \in \mathscr{U}$

$$\begin{aligned} &\langle u, (J(s) + [J(s)]^*)u\rangle_{\mathscr{U}} \\ &\quad = 2\langle (I+G(s))^{-1}u, (I - (G(s))^*G(s))(I+G(s))^{-1}u\rangle_{\mathscr{U}} \\ &\quad \geq 0, \end{aligned} \tag{7.11}$$

by the inequality (7.9). We conclude that $J \in PR$. Conversely, if $J \in PR \subseteq D(\tilde{\mathcal{S}})$ then setting $G := \tilde{\mathcal{S}}(J)$ gives

$$\begin{aligned} I - G^*G &= I - (I+J)^{-*}(I-J)^*(I-J)(I+J)^{-1} \\ &= (I+J)^{-*}\left[(I+J)^*(I+J) - (I-J)^*(I-J)\right](I+J)^{-1} \\ &= 2(I+J)^{-*}[J + J^*](I+J)^{-1}. \end{aligned}$$

Hence for $s \in \mathbb{C}_0^+$ and $u \in \mathscr{U}$

$$\begin{aligned} &\langle u, (I - [G(s)]^*G(s))u\rangle_{\mathscr{U}} \\ &\quad = 2\langle (I+J(s))^{-1}u, ((J(s))^* + J(s))(I+J(s))^{-1}u\rangle_{\mathscr{U}} \\ &\quad \geq 0, \end{aligned} \tag{7.12}$$

as $J \in PR$. The characterisation (7.9) above of bounded realness implies that $G \in BR$ and as $\tilde{\mathcal{S}} : D(\tilde{\mathcal{S}}) \to D(\tilde{\mathcal{S}})$ we infer that $G \in BR \cap D(\tilde{\mathcal{S}})$. We have proven that

$$\begin{aligned} &\tilde{\mathcal{S}}(BR \cap D(\tilde{\mathcal{S}})) \subseteq PR, \\ &\quad \text{and} \quad \tilde{\mathcal{S}}(PR) \subseteq BR \cap D(\tilde{\mathcal{S}}), \end{aligned}$$

and since $\tilde{\mathcal{S}}$ is its own inverse on $D(\tilde{\mathcal{S}})$ we conclude that (v) holds.

(vi): Choose $G \in SBR \subseteq D(\tilde{\mathcal{S}})$ (by (ii)) then from (7.11) for $J := \tilde{\mathcal{S}}(G)$ we see that

$$\begin{aligned}\langle u,(J(s)+J^*(s))u\rangle_{\mathscr{U}} &= 2\langle (I+G(s))^{-1}u,(I-(G(s))^*G(s))(I+G(s))^{-1}u\rangle_{\mathscr{U}} \\ &\geq \varepsilon\|(I+G(s))^{-1}u\|_{\mathscr{U}} \geq \varepsilon c_2\|u\|_{\mathscr{U}},\end{aligned}$$

by (7.10) and (7.4). Thus $J=\tilde{\mathcal{S}}(G)\in SPR$, which is (vi).

(vii): For $J\in SPR\cap H^\infty$ and $G:=\tilde{\mathcal{S}}(J)$, inserting (2.23) and (7.8) into (7.12) yields for $s\in\mathbb{C}_0^+$ and $u\in\mathscr{U}$

$$\begin{aligned}\langle u,(I-[G(s)]^*G(s))u\rangle_{\mathscr{U}} &= 2\langle (I+J(s))^{-1}u,((J(s))^*+J(s))(I+J(s))^{-1}u\rangle_{\mathscr{U}} \\ &\geq \eta\|(I+J(s))^{-1}u\|_{\mathscr{U}} \geq \eta c_4\|u\|_{\mathscr{U}}, \quad c_4,\eta>0.\end{aligned}$$

We conclude that $G\in SBR\subseteq H^\infty$. Summarising, we have proven that

$$\begin{aligned}\tilde{\mathcal{S}}(SBR)\subseteq SPR,\ \tilde{\mathcal{S}}(SBR)\subseteq H^\infty \quad &\Rightarrow \quad \tilde{\mathcal{S}}(SBR)\subseteq SPR\cap H^\infty \\ &\text{and} \quad \tilde{\mathcal{S}}(SPR\cap H^\infty)\subseteq SBR,\end{aligned}$$

and since $\tilde{\mathcal{S}}$ is its own inverse on $D(\tilde{\mathcal{S}})$, we infer that (vii) holds completing the proof. □

Remark 7.1.9. Note that

$$\tilde{\mathcal{S}}:BR\to PR,$$

is not bijective (it isn't even well-defined) as Example 7.1.5 demonstrates that there exist BR functions which do not belong to $D(\tilde{\mathcal{S}})$. Secondly,

$$\tilde{\mathcal{S}}:SBR\to SPR,$$

is also not bijective as the image under $\tilde{\mathcal{S}}$ of every SBR function belongs to H^∞ (part (iv) above), but there are functions in SPR that are not in H^∞, for example $J(s)\equiv s+d$, $d\in\mathbb{R}$, $d>0$. Claim (vii) above could be rephrased as

$$\tilde{\mathcal{S}}:SBR\cap H^\infty\to SPR\cap H^\infty,$$

is bijective, but that $SBR\subseteq H^\infty$.

Corollary 7.1.10. *For $\mathscr{U}$ finite-dimensional let $G\in H^\infty(\mathbb{C}_0^+;B(\mathscr{U}))$ be strictly bounded real. Then G is regular if and only if $\tilde{\mathcal{S}}(G)$ is.*

Proof. Since $\tilde{\mathcal{S}}$ is self-inverse it suffices to prove just one direction. Assume that G is regular with feedthrough D which satisfies

$$D=\lim_{\substack{s\to\infty\\ s\in\mathbb{R}^+}}G(s) \quad\Rightarrow\quad I+D=\lim_{\substack{s\to\infty\\ s\in\mathbb{R}^+}}I+G(s).$$

By Lemma 7.1.7, $I + G(s)$ is boundedly invertible, uniformly in $s \in \mathbb{C}_0^+$, and hence so is the above limit. We conclude that

$$s \mapsto (I + G(s))^{-1},$$

is also regular. Therefore the product

$$s \mapsto (I - G(s))(I + G(s))^{-1} = \tilde{\mathcal{S}}(G)(s),$$

is regular by the algebra of limits. $\square$

The next result is contained within [79, Theorem 5.2], although the formulae (7.14) are not given there, and demonstrates that given a well-posed realisation of a strictly bounded real function G, we can obtain a well-posed realisation of (the strictly positive real function) $\tilde{\mathcal{S}}(G)$ with the same state.

Lemma 7.1.11. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ be strictly bounded real. If $\Sigma_G = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ on $(\mathscr{U}, \mathscr{X}, \mathscr{U})$ is an L^p $(1 \leq p < \infty)$ well-posed realisation of G and*

$$v = \frac{u + y}{\sqrt{2}}, \quad w = \frac{u - y}{\sqrt{2}}, \tag{7.13}$$

then

$$\begin{aligned} \Sigma_{\tilde{\mathcal{S}}(G)} &= \begin{bmatrix} \mathfrak{A} - \mathfrak{B}(I + \mathfrak{D})^{-1}\mathfrak{C} & \sqrt{2}\mathfrak{B}(I + \mathfrak{D})^{-1} \\ -\sqrt{2}(I + \mathfrak{D})^{-1}\mathfrak{C} & (I - \mathfrak{D})(I + \mathfrak{D})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathfrak{A} - \mathfrak{B}(I + \mathfrak{D})^{-1}\mathfrak{C} & \sqrt{2}\mathfrak{B}(I + \mathfrak{D})^{-1} \\ -\sqrt{2}(I + \mathfrak{D})^{-1}\mathfrak{C} & \check{\mathcal{S}}(\mathfrak{D}) \end{bmatrix}, \end{aligned} \tag{7.14}$$

is an L^p well-posed realisation of $\tilde{\mathcal{S}}(G)$ on $(\mathscr{U}, \mathscr{X}, \mathscr{U})$. Moreover the state trajectories of Σ_G with input u and output y and $\Sigma_{\tilde{\mathcal{S}}(G)}$ with input v and output w are the same.

Proof. See [79, Theorem 5.2]. As mentioned in the proof of that result, the relationship

$$v = \frac{u + y}{\sqrt{2}} \quad \Rightarrow \quad u = \sqrt{2}v - y,$$

can be seen as (negative identity) static output feedback with external control v. The relationship

$$w = \frac{u - y}{\sqrt{2}} \quad \Rightarrow \quad w = v - \sqrt{2}y,$$

corresponds to adding an extra feedthrough term. From these observations and the formulae for the closed loop well-posed linear system from [81, Theorem 7.1.2] the formulae in (7.14) follow. $\square$

Remark 7.1.12. The above result also has a natural converse. Given an L^p ($1 \leq p < \infty$) well-posed realisation $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ on $(\mathscr{U}, \mathscr{X}, \mathscr{U})$ of a strictly positive real $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ then the realisation in (7.14) is a L^p well-posed realisation of $\tilde{S}(J)$. The proof is exactly the same.

7.2 Positive real balanced truncation

In this section we define the positive real balanced truncation of a strictly positive real transfer function with summable positive real singular values and prove a gap metric error bound, formulated as Theorem 7.2.12. To do so we make use of the material gathered in Section 7.1.

We are now considering a strictly positive real function $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, as in Definition 2.3.1. Our first aim is to define the positive real singular values. Since $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, by Lemma 4.1.5 there exist stable L^2 well-posed realisations of J. The positive real optimal control problem is formulated and solved as in the bounded real case.

Lemma 7.2.1. *Let $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ denote a stable L^2 well-posed linear system and assume that Σ has strictly positive real transfer function $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$. Then the optimal control problem: for $x_0 \in \mathscr{X}$ minimise*

$$\mathcal{L}(x_0, u) = \int_{\mathbb{R}^+} 2\mathrm{Re}\, \langle u(s), y(s) \rangle_{\mathscr{U}} \, ds, \tag{7.15}$$

over all $u \in L^2(\mathbb{R}^+; \mathscr{U})$ subject to (4.1), *has a solution in the sense that for any $x_0 \in \mathscr{X}$*

$$\inf_{u \in L^2(\mathbb{R}^+; \mathscr{U})} \mathcal{L}(x_0, u) = \mathcal{L}(x_0, \tilde{u}_{opt}) = -\langle \tilde{P}_m x_0, x_0 \rangle_{\mathscr{X}}. \tag{7.16}$$

The optimal control is uniquely given by

$$\tilde{u}_{opt} = -(\mathfrak{D}\pi_+ + \pi_+\mathfrak{D}^*)^{-1} \mathfrak{C}\, x_0, \tag{7.17}$$

and $\tilde{P}_m : \mathscr{X} \to \mathscr{X}$ is bounded and satisfies $P_m = P_m^* \geq 0$ and

$$\tilde{P}_m = \mathfrak{C}^* (\mathfrak{D}\pi_+ + \pi_+\mathfrak{D}^*)^{-1} \mathfrak{C}. \tag{7.18}$$

Proof. See [97, Proposition 7.2]. Note that the assumption that J is strictly positive real is equivalent to

$$\mathfrak{D}\pi_+ + \pi_+\mathfrak{D}^* \geq \varepsilon I,$$

see [97, Section 7] and hence $(\mathfrak{D}\pi_+ + \pi_+\mathfrak{D}^*)$ is boundedly invertible. Therefore the optimal control $\tilde{u}_{opt}$ and optimal cost operator $\tilde{P}_m$ in (7.17) and (7.18) respectively are well-defined. Furthermore, in [97] it is assumed that J is weakly regular (with zero feedthrough), but that is not needed for this proof. $\square$

For $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ a strictly positive real transfer function, we let J^d denote the dual transfer function as in Definition 4.2.1. Again it is easy to see that J is (strictly) positive real if and only J_d is. Given a stable L^2 well-posed realisation of J, we let $\tilde{Q}_m : \mathscr{X} \to \mathscr{X}$ denote the self-adjoint, non-negative optimal cost operator for the dual positive real optimal control problem, formulated analogously to the dual bounded real optimal control problem in Lemma 6.2.2.

The next result is crucial for linking positive real balanced truncation to bounded real balanced truncation.

Lemma 7.2.2. *Let Σ_J denote a stable L^2 well-posed linear system and assume that Σ_J has strictly positive real transfer function $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$. Let $\tilde{P}_m$ and $\tilde{Q}_m$ denote the optimal cost operators of the positive real optimal control problem* (7.15) *and the dual problem subject to Σ_J respectively. Let $\Sigma_{\tilde{\mathcal{S}}(J)}$ denote the L^2 well-posed realisation given by* (7.14). *Then the optimal cost operators of the bounded real optimal control problem* (6.1) *and the dual problem* (6.5) *subject to $\Sigma_{\tilde{\mathcal{S}}(J)}$ are $\tilde{P}_m$ and $\tilde{Q}_m$ respectively.*

Proof. Let $\Sigma_J = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ so that by equation (7.14) the output map and input-output map of $\Sigma_{\tilde{\mathcal{S}}(J)}$ are given by

$$\mathfrak{C}_{\tilde{\mathcal{S}}(J)} = -\sqrt{2}(I + \mathfrak{D})^{-1}\mathfrak{C}, \quad \check{\mathcal{S}}(\mathfrak{D}) = (I - \mathfrak{D})(I + \mathfrak{D})^{-1}, \tag{7.19}$$

respectively. Let P_m denote the optimal cost of bounded real optimal control problem (6.1) subject to the realisation $\Sigma_{\tilde{\mathcal{S}}(J)}$. A long, but elementary, calculation using (7.19) shows that

$$\begin{aligned}
\tilde{P}_m &= \mathfrak{C}^*((\mathfrak{D}\pi_+ + \pi_+\mathfrak{D}^*))^{-1}\mathfrak{C}, \quad \text{from (7.18)},\\
&= \mathfrak{C}^*_{\tilde{\mathcal{S}}(J)}\mathfrak{C}_{\tilde{\mathcal{S}}(J)} + \mathfrak{C}^*_{\tilde{\mathcal{S}}(J)}\check{\mathcal{S}}(\mathfrak{D})\pi_+(I - \pi_+\check{\mathcal{S}}(\mathfrak{D})^*\check{\mathcal{S}}(\mathfrak{D})\pi_+)^{-1}\pi_+\check{\mathcal{S}}(\mathfrak{D})^*\mathfrak{C}_{\tilde{\mathcal{S}}(J)}\\
&= P_m, \quad \text{from (6.4)},
\end{aligned}$$

as required. The dual argument is exactly the same, using instead the dual L^2 well-posed linear systems, which are also related by Lemma 7.1.11. $\square$

Definition 7.2.3. Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly positive real transfer function. We define the positive real singular values of J as the bounded real singular values of the strictly bounded real function $G := \tilde{\mathcal{S}}(J)$.

The next result shows that the above definition of positive real singular values is consistent with the definition of positive real singular values in the finite-dimensional case, described in Section 2.3.2. We prove that when the positive real singular values are summable (with their multiplicities) then they are the square roots of the eigenvalues of the product of the optimal cost operators for the positive real optimal control problems, independently of the realisation chosen.

Lemma 7.2.4. *Let $J \in H^\infty(\mathbb{C}_0^+; B(U))$ be strictly positive real and let $\Sigma = (\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D})$ denote a stable L^2 well-posed linear system realising J. Let $\tilde{P}_m$ and $\tilde{Q}_m$ denote the optimal cost operators of the positive real optimal control problem* (7.15) *and the dual positive real optimal control problem subject to Σ respectively. Then the positive real singular values are summable if and only if $\tilde{P}_m\tilde{Q}_m$ is compact and the square roots of its eigenvalues are summable. If these conditions hold then apart from possibly zero the positive real singular values are precisely the square roots of the eigenvalues of $\tilde{P}_m\tilde{Q}_m$ (which depend only on J and as such are independent of the stable L^2 well-posed linear system realising J).*

Proof. This follows immediately from the definition of positive real singular values, Lemma 6.3.12 and Lemma 7.2.2. $\square$

Corollary 7.2.5. *If $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ is strictly positive real with summable positive real singular values, then J is regular.*

Proof. Set $G := \tilde{\mathcal{S}}(J)$, which by Lemma 7.1.8 is strictly bounded real and has summable bounded real singular values by Definition 7.2.3. From Lemma 6.3.7 it follows that G is regular, and hence so is J by Corollary 7.1.10. $\square$

The next lemma prepares the positive real balanced truncation of strictly positive real functions with summable positive real singular values. We obtain a family of L^1 well-posed realisations of J, using the Cayley transform, that we will truncate in Definition 7.2.8 to give a family of positive real balanced truncations.

Lemma 7.2.6. *Given $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ a strictly positive real transfer function with summable positive real singular values, set $G := \tilde{\mathcal{S}}(J)$, which is strictly bounded real and has summable bounded real singular values. Let G_E denote a member of the family of extended transfer functions of G and let (A_E, B_E, C_E, D_E) denote the generators of the exactly observable shift realisation on L^1 of G_E. Let A, B, C and D denote the generators of the L^1 well-posed realisation of G obtained from (A_E, B_E, C_E, D_E) by Lemma 6.3.9. The operators*

$$\begin{aligned} \tilde{A} &= A - B(I+D)^{-1}C : D(A) \to \mathscr{X}, & \tilde{B} &= \sqrt{2}B(I+D)^{-1} : \mathscr{U} \to \mathscr{X}, \\ \tilde{C} &= -\sqrt{2}(I+D)^{-1}C : D(A) \to \mathscr{U}, & \tilde{D} &= (I-D)(I+D)^{-1} : \mathscr{U} \to \mathscr{U}, \end{aligned} \tag{7.20}$$

are well-defined and are the generators of an L^1 well-posed realisation for J. In particular,

$$J(s) = \tilde{D} + \tilde{C}(sI - \tilde{A})^{-1}\tilde{B}, \quad s \in \mathbb{C}_0^+. \tag{7.21}$$

Proof. The function G is strictly bounded real by Lemma 7.1.8, and has summable bounded real singular values by Definition 7.2.3. Therefore, we can choose an extended transfer function G_E, exactly observable shift realisation on L^1 of G_E and the resulting generators of an L^1 well-posed realisation of G according to the statement of the lemma.

We transform the L^1 well-posed realisation of G generated by A, B, C and D as in Lemma 7.1.11, to give an L^1 well-posed realisation of J. The generators $\tilde{A}, \tilde{B}, \tilde{C}$ and $\tilde{D}$ of this realisation are given by [81, Theorem 7.5.1 (ii)] and [81, Lemma 5.1.2 (ii)], where we have used the boundedness of B to infer that A, B, C and D generate a compatible system node with $W = D(A)$. Note that there are changes from our (7.14) and [81, (7.1.5)] because we combined a feedback with an extra feedthrough term. As such the generators have also changed accordingly. The formula (7.21) follows from [81, Theorem 4.6.3 (ii)]. $\square$

Remark 7.2.7. The result of Lemma 7.2.6 is an infinite-dimensional version of [57, Lemma 3]. We remark, however, that the transformation (15) in [57] is *not* the same transformation as (7.1.3). As such the formulae in (7.20) are slightly different to those in [57, Lemma 3]; namely there is a difference in signs.

Definition 7.2.8. Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function with summable positive real singular values, and let G_E denote a member of the family of extended transfer functions of $G := \tilde{\mathcal{S}}(J)$. Let $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ denote the generators of the L^1 well-posed realisation of J from Lemma 7.2.6. We define the operators $\tilde{A}_n$, $\tilde{B}_n$ and $\tilde{C}_n$ by

$$\begin{aligned} &\tilde{A}_n := \mathcal{P}_{\mathscr{X}_n} \tilde{A}|_{\mathscr{X}_n} : \mathscr{X}_n \to \mathscr{X}_n, \quad \tilde{B}_n := \mathcal{P}_{\mathscr{X}_n} \tilde{B} : \mathscr{U} \to \mathscr{X}_n, \\ &\tilde{C}_n := \tilde{C}|_{\mathscr{X}_n} : \mathscr{X}_n \to \mathscr{U}, \end{aligned} \tag{7.22}$$

where $\mathscr{X}_n$ is the truncation space (5.72). The input-state-output system generated by $\begin{bmatrix} \tilde{A}_n & \tilde{B}_n \\ \tilde{C}_n & \tilde{D} \end{bmatrix}$ is called the reduced order system obtained by positive real balanced truncation (determined by G_E). We call J_n given by

$$J_n(s) := \tilde{C}_n(sI - \tilde{A}_n)^{-1}\tilde{B}_n + \tilde{D},$$

defined on some right half-plane, the reduced order transfer function obtained by positive real balanced truncation.

The next lemma demonstrates that the positive real balanced truncation is determined by J up to a unitary transformation and thus that the reduced order transfer function J_n is uniquely determined by J.

Lemma 7.2.9. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function with summable positive real singular values, and let G_E denote a member of the family* *of extended transfer functions of* $G := \tilde{\mathcal{S}}(J)$*. For* $n \in \mathbb{N}$ *let* $\left[\begin{smallmatrix} A_n & B_n \\ C_n & D \end{smallmatrix}\right]$ *and* $\left[\begin{smallmatrix} \tilde{A}_n & \tilde{B}_n \\ \tilde{C}_n & \tilde{D} \end{smallmatrix}\right]$ *denote the bounded real and positive real balanced truncations (determined by* G_E*) of* G *and* J *respectively, with respective transfer functions* G_n *and* J_n*. Then*

(i) We have the following relations between the positive real and bounded real balanced truncations

$$\begin{aligned} \tilde{A}_n &= A_n - B_n(I+D)^{-1}C_n, & \tilde{B}_n &= \sqrt{2}B_n(I+D)^{-1}, \\ \tilde{C}_n &= -\sqrt{2}(I+D)^{-1}C_n, & \tilde{D} &= (I-D)(I+D)^{-1}. \end{aligned} \tag{7.23}$$

(ii) J_n *is proper rational and positive real.*

(iii) Different choices of G_E *gives rise to positive real balanced truncations that are unitarily equivalent, so that every choice of* G_E *gives rise to the same* J_n *and the following commutative diagram holds*

$$\begin{array}{ccc} J & \xrightarrow{\tilde{\mathcal{S}}} & \tilde{\mathcal{S}}(J) \\ {\scriptstyle\text{prbt}}\big\downarrow & & {\scriptstyle\text{brbt}}\big\downarrow \\ J_n & \xrightarrow{\tilde{\mathcal{S}}} & \tilde{\mathcal{S}}(J)_n \end{array}$$

As such, $G_n \in D(\tilde{\mathcal{S}})$ *and* $J_n = \tilde{\mathcal{S}}(G_n)$*.*

Proof. That (7.23) holds follows from the definition of $\left[\begin{smallmatrix} A_n & B_n \\ C_n & D \end{smallmatrix}\right]$ in Definition 6.3.13, that of $\left[\begin{smallmatrix} \tilde{A}_n & \tilde{B}_n \\ \tilde{C}_n & \tilde{D} \end{smallmatrix}\right]$ in Definition 7.2.8 and the fact that restriction and projection are linear operations. That different choices of G_E give rise to unitarily equivalent positive real balanced truncations now follows from the relations (7.23) and Lemma 6.3.14. In particular, every choice of G_E gives rise to the same reduced order transfer function J_n obtained by positive real balanced truncation.

An elementary, but tedious, calculation demonstrates that if (P, K, W) solve the bounded real Lur'e equations (2.10) subject to the realisation $\left[\begin{smallmatrix} A_n & B_n \\ C_n & D \end{smallmatrix}\right]$ then (P, K', W') solve the positive real Lur'e equations (2.24) subject to $\left[\begin{smallmatrix} \tilde{A}_n & \tilde{B}_n \\ \tilde{C}_n & \tilde{D} \end{smallmatrix}\right]$ where

$$K' = K - W(I+D)^{-1}C_n, \quad W' = \sqrt{2}W(I+D)^{-1}.$$

From the Positive Real Lemma it follows that J_n is positive real and it is clearly rational since it has a realisation with finite-dimensional state-space. Therefore by Lemma 7.1.8 (*ii*), $J_n \in D(\tilde{\mathcal{S}})$ and another elementary calculation using (7.23) shows that $\tilde{\mathcal{S}}(J_n) = G_n$. Therefore by Lemma 7.1.8 (*v*), $G_n \in D(\tilde{\mathcal{S}})$ and $\tilde{\mathcal{S}}(G_n) = J_n$.

We note that the commutative diagram is well defined in the sense that it is independent of G_E. Furthermore, the above observations have demonstrated that it does indeed commute. ☐

We now gather some lemmas required to prove our main result of the chapter; a gap metric error bound for positive real balanced truncation which is formulated as Theorem 7.2.12 below. The gap between closed operators was recalled in Definition 3.1.28.

Lemma 7.2.10. *For $\mathscr{U}$ a finite-dimensional Hilbert space let F denote the map*

$$F : \begin{bmatrix} L^2(\mathbb{R}^+;\mathscr{U}) \\ L^2(\mathbb{R}^+;\mathscr{U}) \end{bmatrix} \to \begin{bmatrix} L^2(\mathbb{R}^+;\mathscr{U}) \\ L^2(\mathbb{R}^+;\mathscr{U}) \end{bmatrix}, \quad F = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ I & -I \end{bmatrix}, \tag{7.24}$$

which is an isometric isomorphism mapping pairs $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} v \\ w \end{smallmatrix}\right]$, where u, y, v, w satisfy (7.13)*. Given the Cayley transform $\check{\mathcal{S}}$ of Remark 7.1.2 and $\mathfrak{D} \in D(\check{\mathcal{S}})$, it follows that*

$$F\mathcal{G}(\mathfrak{D}) = \mathcal{G}(\check{\mathcal{S}}(\mathfrak{D})),$$

where $\mathcal{G}(\mathfrak{D})$ denotes the graph of $\mathfrak{D}$.

Proof. The simple proof is left to the reader. □

The following elementary lemma shows that the gap metric is invariant under isometries.

Lemma 7.2.11. *For $\mathscr{M}, \mathscr{N} \subseteq \mathscr{X}$ closed subspaces of a Hilbert space $\mathscr{X}$ and $T : \mathscr{X} \to \mathscr{X}$ an isometry we have*

$$\hat{\delta}(T\mathscr{M}, T\mathscr{N}) = \hat{\delta}(\mathscr{M}, \mathscr{N}).$$

Proof.

$$\delta(T\mathscr{M}, T\mathscr{N}) = \sup_{\substack{m \in T\mathscr{M} \\ \|m\|=1}} \inf_{n \in T\mathscr{N}} \|m - n\| = \sup_{\substack{m' \in \mathscr{M} \\ \|Tm'\|=1}} \inf_{n' \in \mathscr{N}} \|Tm' - Tn'\|.$$

Now using that T is an isometry

$$\delta(T\mathscr{M}, T\mathscr{N}) = \sup_{\substack{m' \in \mathscr{M} \\ \|m'\|=1}} \inf_{n' \in \mathscr{N}} \|m' - n'\| = \delta(\mathscr{M}, \mathscr{N}).$$

The same argument applies for $\delta(T\mathscr{N}, T\mathscr{M})$ and hence

$$\begin{aligned} \hat{\delta}(T\mathscr{M}, T\mathscr{N}) &= \max[\delta(T\mathscr{M}, T\mathscr{N}), \delta(T\mathscr{N}, T\mathscr{M})] \\ &= \max[\delta(\mathscr{M}, \mathscr{N}), \delta(\mathscr{N}, \mathscr{M})] = \hat{\delta}(\mathscr{M}, \mathscr{N}), \end{aligned}$$

as required. □

Theorem 7.2.12. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function with summable positive real singular values $(\sigma_k)_{k\in\mathbb{N}}$ and where $\mathscr{U}$ is finite-dimensional. Then for each integer n, the reduced order transfer function J_n obtained by positive real balanced truncation from Definition 7.2.8 is positive real, proper rational and satisfies the bound*

$$\hat{\delta}(J, J_n) \leq 2 \sum_{k\geq n+1} \sigma_k. \tag{7.25}$$

In inequality (7.25) *we are abusing notation by writing $\hat{\delta}(J, J_r) = \hat{\delta}(\mathfrak{D}_J, \mathfrak{D}_{J_r})$, where $\mathfrak{D}_J$ and $\mathfrak{D}_{J_r}$ are the input-output maps corresponding to J and J_r respectively.*

Proof. Since J is strictly positive real with summable bounded real singular values, the hypotheses of Lemma 7.2.6 are satisfied and thus the positive real balanced truncation J_n of Definition 7.2.8 is well-defined. That J_n is proper rational and positive real now follows from Lemma 7.2.9 (*ii*). It remains to prove the error bound (7.25). From Lemmas 7.1.8 and 7.2.4 the transfer function $G := \tilde{\mathcal{S}}(J)$ is strictly bounded real with summable bounded real singular values. Therefore all the assumptions of Theorem 6.3.15 are satisfied and so the error bound (6.47) holds for G and its bounded real balanced truncation G_n.

Let $\mathfrak{D}_G$ and $\mathfrak{D}_{G_n}$ denote the input-output maps of G and G_n respectively. From the commuting diagram in Lemma 7.2.9 it follows that $\check{\mathcal{S}}(\mathfrak{D}_{G_n}) = \mathfrak{D}_{J_n}$. Therefore we compute

$$\begin{aligned}\hat{\delta}(\mathcal{G}(\mathfrak{D}_J), \mathcal{G}(\mathfrak{D}_{J_n})) &= \hat{\delta}(\mathcal{G}(\check{\mathcal{S}}(\mathfrak{D}_G)), \mathcal{G}(\check{\mathcal{S}}(\mathfrak{D}_{G_n}))) \\ &= \hat{\delta}(F\mathcal{G}(\mathfrak{D}_G), F\mathcal{G}(\mathfrak{D}_{G_n})), \quad \text{by Lemma 7.2.10,} \\ &= \hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_n})), \quad \text{by Lemma 7.2.11.}\end{aligned} \tag{7.26}$$

The bound (3.39) from Theorem 3.1.32 gives

$$\hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_n})) \leq \|\mathfrak{D}_G - \mathfrak{D}_{G_n}\|, \tag{7.27}$$

and it is well-known that

$$\|\mathfrak{D}_G - \mathfrak{D}_{G_n}\| = \|G - G_n\|_{H^\infty}, \tag{7.28}$$

(see for example [94]). Combining (7.26), (7.27), (7.28) and (6.47) yields

$$\begin{aligned}\hat{\delta}(\mathcal{G}(\mathfrak{D}_J), \mathcal{G}(\mathfrak{D}_{J_n})) = \hat{\delta}(\mathcal{G}(\mathfrak{D}_G), \mathcal{G}(\mathfrak{D}_{G_n})) &\leq \|\mathfrak{D}_G - \mathfrak{D}_{G_n}\| \\ &= \|G - G_n\|_{H^\infty} \leq 2 \sum_{k\geq n+1} \sigma_k,\end{aligned}$$

which is (7.25). Finally we note that $(\sigma_k)_{k\in\mathbb{N}}$ are the bounded real singular values of

G which by definition are the positive real singular values of J. $\square$

7.3 Notes

The aim of this chapter is to extend model reduction by positive real balanced truncation to infinite-dimensional systems. The main result of this chapter is Theorem 7.2.12, which is new. A version of this chapter, together with material from Chapters 6 and 8, has been submitted for publication as [37]. As with the finite-dimensional case, positive real balanced truncation can be derived from bounded real balanced truncation via the Cayley transform, which has been our approach here. The material on the Cayley transform is not new and is a common tool in network analysis, for example as in [2] as a transform from an impedance matrix to a scattering matrix. We could not find explicit statements of results such as Lemma 7.1.8 in the literature, although the conclusions of that lemma must be known.

Chapter 8

Remarks on applications and an example

Our main results for infinite-dimensional balanced truncation, Theorems 6.3.15 and 7.2.12, each have two key assumptions. We require that the transfer function is strictly bounded real (respectively strictly positive real) and has summable bounded real singular values (respectively summable positive real singular values). In this penultimate chapter we comment on these assumptions and provide an example. Recall from Chapter 6 that one of our motivations for extending model reduction by balanced truncation to the infinite-dimensional case (especially when compared to existing model reduction schemes) was the much faster rate of convergence of the truncations. In Section 8.1 we consider a class of bounded real and positive real systems where the corresponding singular values decay very rapidly (and in particular are summable). In Section 8.2 we consider an example of a controlled partial differential equation and demonstrate the results of a numerical calculation of approximate positive real balanced truncations.

8.1 Asymptotic behavior of bounded real and positive real singular values

Theorem 6.3.15 requires that the transfer function G has summable bounded real singular values. As we have seen in Lemma 6.3.12, the bounded real singular values are summable precisely when the Hankel singular values of a (equivalently every) member of the family of extended Hankel operators of G are summable. Therefore, we seek conditions which ensure that a Hankel operator is nuclear. The next result is taken from [61]. In what follows $\mathscr{X}_\alpha$ denote interpolation spaces, see for example, [81, Section 3.9] and S_p is the Schatten class from Definition 4.3.2.

Theorem 8.1.1. *Assume that A generates an exponentially stable analytic semigroup, $B \in B(\mathscr{U}, \mathscr{X}_\beta)$, $C \in B(\mathscr{X}_\alpha, \mathscr{Y})$ and $D \in B(\mathscr{U}, \mathscr{Y})$, with $\alpha - \beta < 1$ and that at least* *one of $\mathscr{U}$ and $\mathscr{Y}$ is finite-dimensional. Then the Hankel operator of this system is in S_p for all $p > 0$.*

Given a stable L^2 well-posed realisation of the strictly bounded real transfer function G with generators (A, B, C, D), and choice of spectral factors θ and ξ as in Lemma 6.2.3 it follows from Lemma 6.3.7 that (A, B_E, C_E, D_E) generate a stable L^2 well-posed realisation of the extended transfer function G_E. Here

$$B_E = \begin{bmatrix} B & B_\xi \end{bmatrix}, \quad C_E = \begin{bmatrix} C \\ C_\theta \end{bmatrix},$$

are the generators of $\mathfrak{B}_E$ and $\mathfrak{C}_E$ from (6.20) and (6.12) respectively and D_E is as in (6.33). It is not *a priori* clear how unbounded C_E and B_E are because it is not presently clear how unbounded the components C_θ and B_ξ are. However, under the assumption of strict bounded realness, we are able to formulate the next result which provides checkable conditions for the summability of the bounded real singular values and in fact in this case ensures a rapid decay rate.

Proposition 8.1.2. *Let $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}, \mathscr{Y}))$ denote a strictly bounded real transfer function and assume that G is generated by (A, B, C, D) where A generates an exponentially stable analytic semigroup, $B \in B(\mathscr{U}, \mathscr{X}_\beta)$, $C \in B(\mathscr{X}_\alpha, \mathscr{Y})$ and $D \in B(\mathscr{U}, \mathscr{Y})$, with $\alpha - \beta < 1$ and that both $\mathscr{U}$ and $\mathscr{Y}$ are finite-dimensional. Then the bounded real singular values of G belong to ℓ^p for every $p > 0$. In particular, they are summable and moreover decay faster than any polynomial rate, that is for any $p > 0$*

$$n^p \sigma_n \to 0, \quad \textit{as} \quad n \to \infty.$$

Proof. In Staffans [76, Theorem 1] it is proven that under our assumptions the operator C_θ from (6.51) is bounded $\mathscr{X}_\alpha \to \mathscr{U}$. Hence C_E is bounded $\mathscr{X}_\alpha \to \left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$. Repeating the argument in the dual case (so where G is replaced by G_d, which is strictly bounded real, with realisation $\left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$) we deduce that the generator C_η of $\mathfrak{C}_\eta$ from (6.25) satisfies

$$C_\eta \in B((\mathscr{X}^*)_{-\beta}, \mathscr{Y}),$$

where $(\mathscr{X}^*)_\gamma$ denote the interpolation spaces corresponding to A^*. By construction the generator B_ξ of $\mathfrak{B}_\xi$ in (6.22) is equal to C_η^* Since $\mathscr{X}$ is a Hilbert space, in the above we can identify $(\mathscr{X}^*)_\gamma$ with $\mathscr{X}_{-\gamma}$ and hence

$$B_\xi = C_\eta^* \in B(\mathscr{Y}, (\mathscr{X}^*)_{-\beta}) = B(\mathscr{Y}, \mathscr{X}_\beta).$$

We conclude that B_E is bounded $\left[\begin{smallmatrix} \mathscr{U} \\ \mathscr{Y} \end{smallmatrix}\right] \to \mathscr{X}_\beta$. Therefore from Theorem 8.1.1, the Hankel operator of the extended system belongs to S_p and hence the bounded real singular values belong to ℓ^p. $\square$

The next result is a corresponding version of the above for positive real systems.

Corollary 8.1.3. *Let $J \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$ denote a strictly positive real transfer function, with $\mathscr{U}$ finite-dimensional. Assume that J is generated by (A, B, C, D) where A generates an exponentially stable analytic semigroup on $\mathscr{X}$, $B \in B(\mathscr{U}, \mathscr{X}_\beta)$, $C \in B(\mathscr{X}_\alpha, \mathscr{Y})$ and $D \in B(\mathscr{U}, \mathscr{Y})$, with $\alpha \in [0,1]$ and $\alpha - \beta < 1$. Then the positive real singular values of J belong to ℓ^p for every $p > 0$. In particular, they are summable and moreover decay faster than any polynomial rate, that is for any $p > 0$*

$$n^p \sigma_n \to 0, \quad \textit{as} \quad n \to \infty.$$

Proof. From Lemma 7.1.8, the function $G := \tilde{S}(J)$ is strictly bounded real and from Lemma 7.2.4 the bounded real singular values of G are the positive real singular values of J. We seek therefore to apply Proposition 8.1.2, and in order to do so we require a state-space realisation of G. As argued in the proof of Lemma 7.2.6, the Cayley transform of operators

$$\begin{aligned}
\tilde{A} &= A|_{\mathscr{X}_\alpha} - B(I+D)^{-1}C : \mathscr{X}_\alpha \to \mathscr{X}_{\alpha-1}, & \tilde{B} &= \sqrt{2}B(I+D)^{-1} : \mathscr{U} \to \mathscr{X}_\beta, \\
\tilde{C} &= -\sqrt{2}(I+D)^{-1}C : \mathscr{X}_\alpha \to \mathscr{U}, & \tilde{D} &= (I-D)(I+D)^{-1} : \mathscr{U} \to \mathscr{U},
\end{aligned}$$

is well-defined and $\left[\begin{smallmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{smallmatrix}\right]$ is a realisation of G. This follows again from [81, Theorem 7.5.1 (ii)], here using that $W = \mathscr{X}_\alpha$ is a compatible extension of $\mathscr{X}_1$ (see also [81, Lemma 5.1.2 (iii)]). From Curtain *et al.* [18, Proposition 4.5] the operator $\tilde{A}$ (where $-B(I+D)^{-1}C = \Delta$ in the notation of [18]) generates an analytic semigroup on $\mathscr{X}$ and the interpolation spaces $\mathscr{X}_\delta$ and $\tilde{\mathscr{X}_\delta}$ corresponding to A and $\tilde{A}$ respectively are equal for all $\delta \in [\alpha - 1, \beta + 1]$.

Thus

$$\tilde{B} \in \mathscr{B}(\mathscr{U}, \mathscr{X}_\beta) = \mathscr{B}(\mathscr{U}, \tilde{\mathscr{X}_\beta}), \quad \text{and} \quad \tilde{C} \in \mathscr{B}(\mathscr{X}_\alpha, \mathscr{U}) = \mathscr{B}(\tilde{\mathscr{X}_\alpha}, \mathscr{U}),$$

since trivially $\alpha, \beta \in [\alpha - 1, \beta + 1]$. It remains to see that $\tilde{A}$ generates an exponentially stable semigroup. By the same results from [81] above we can "go back again", and recover the realisation for J from that of G, namely

$$\begin{aligned}
A|_{\mathscr{X}_\alpha} &= \tilde{A} - \tilde{B}(I+\tilde{D})^{-1}\tilde{C} : \mathscr{X}_\alpha \to \mathscr{X}_{\alpha-1}, & B &= \sqrt{2}\tilde{B}(I+\tilde{D})^{-1} : \mathscr{U} \to \mathscr{X}_\beta, \\
C &= -\sqrt{2}(I+\tilde{D})^{-1}\tilde{C} : \mathscr{X}_\alpha \to \mathscr{U}, & D &= (I-\tilde{D})(I+\tilde{D})^{-1} : \mathscr{U} \to \mathscr{U}.
\end{aligned}$$

We now see that $\tilde{A}$ is exponentially stabilisable and detectable since

$$\begin{aligned}
A|_{\mathscr{X}_\alpha} &= \tilde{A} + \tilde{B}F_1, \quad F_1 = -(I+\tilde{D})^{-1}\tilde{C}, \\
A|_{\mathscr{X}_\alpha} &= \tilde{A} + F_2\tilde{C}, \quad F_2 = -\tilde{B}(I+\tilde{D})^{-1},
\end{aligned}$$

and A is exponentially stable. The system with generators $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ is input-output stable, since the transfer function $G \in H^\infty(\mathbb{C}_0^+; B(\mathscr{U}))$, and so by Rebarber [68, Corollary 1.8], $\tilde{A}$ generates an exponentially stable semigroup.

All the hypotheses of Proposition 8.1.2 are satisfied for the realisation $\left[\begin{smallmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{smallmatrix}\right]$ of G, and thus the bounded real singular values of G are in ℓ^p for all $p > 0$. Since the bounded real singular values of G and the positive real singular values of J are the same, this completes the proof. $\square$

8.2 Strictly positive real and strictly bounded real systems: an example

Recall the heat equation system in 1D on the unit interval, formulated as (5.3) in Example 5.0.1. The transfer function J is given by (5.4) which is bounded and analytic in the open right-half plane and so $J \in H^\infty(\mathbb{C}_0^+)$.

From [61] it follows that (5.3) can be written in the form (2.3), with A generating an analytic, exponentially stable semigroup on $\mathscr{X} = L^2(0,1)$. Here C is the trace operator, which is bounded $\mathscr{X}_\alpha \to \mathbb{C}$ for all $\alpha > \frac{1}{4}$. Furthermore, $B = C^*$, and hence B is bounded $\mathbb{C} \to \mathscr{X}_\beta$ for all $\beta < -\frac{1}{4}$. Therefore the conditions on the operators in Corollary 8.1.3 are satisfied and it remains to check the strict positive realness of (5.3). An elementary calculation with

$$E(w(t)) = \int_0^1 \|w(t,x)\|^2 \, dx \geq 0, \quad t \geq 0,$$

gives for smooth solutions w of (5.3)

$$2\mathrm{Re}\,\langle u(t), y(t)\rangle = \frac{d}{dt}E(w(t)) + \underbrace{2\int_0^1 \|w_x\|^2 \, dx}_{\geq 0},$$

$$\Rightarrow \quad \int_0^t 2\mathrm{Re}\,\langle u(\tau), y(\tau)\rangle \, d\tau \geq E(w(\tau)) - E(w(0)), \quad \forall\, t \geq 0. \tag{8.1}$$

Equation (8.1) shows that (5.3) is positive real. However, the system (5.3) is not strictly positive real as $J(s) \to 0$ as $s \to +\infty$ along the real axis and so the condition (2.23) cannot hold.

We can create a strictly positive real transfer function, however, by adding a positive feedthrough to J. Specifically, the system (5.3a)-(5.3d), with (5.3c) replaced by

$$y(t) = w(t,0) + Du(t), \quad D > 0, \tag{5.3e}$$

denoted by $(5.3)'$ is strictly positive real and thus all the hypotheses of Corollary 8.1.3

are satisfied. Hence $(5.3)'$ has summable positive real singular values (belonging to ℓ^p for all $p > 0$, in fact) and positive real balanced truncation is applicable here. Note that the positive real singular values of (5.3) and $(5.3)'$ are not necessarily the same, but both sequences are summable. Note also that in obtaining a strictly positive real transfer function we have changed the boundary conditions of our original PDE.

For illustrative purposes, we also consider the above example in bounded real coordinates. Consider the PDE (5.3a), with boundary condition (5.3d), but now with input u and output y given by

$$u(t) = \frac{1}{\sqrt{2}}(w(t,0) - w_x(t,0)), \tag{5.3f}$$

$$y(t) = -\frac{1}{\sqrt{2}}(w(t,0) + w_x(t,0)). \tag{5.3g}$$

The system (5.3a), (5.3d), (5.3f) and (5.3g) is bounded real, with transfer function

$$G : \mathbb{C}_0^+ \to \mathbb{C}, \quad G(s) = \frac{1 - J(s)}{1 + J(s)} = \frac{\sqrt{s} - \tanh(\sqrt{s})}{\sqrt{s} + \tanh(\sqrt{s})}.$$

The bounded real singular values of G (are the positive real singular values of J and so) are summable. Since $G(0) = 1$, it follows that $\|G\|_{H^\infty} = 1$, and so G is not strictly bounded real. We can alter the system to create a strictly bounded real system by introducing a filter F in series with G as in Fig. 8-1. If the filter F is chosen with

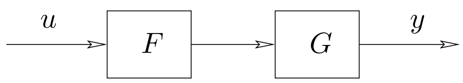


Figure 8-1: *Addition of filter F*

the property that $\|F\|_{H^\infty} < 1$, then the series connection with transfer function GF is strictly bounded real. However, we need to choose F in such a way that the bounded real singular values are still summable. Let $\mathfrak{D}_G$ and $\mathfrak{D}_F$ denote the input-output maps of G and F respectively. The Hankel operator of GF is $\pi_+\mathfrak{D}_G\mathfrak{D}_F\pi_-R$, where R is the reflection operator. A calculation shows that

$$\begin{aligned}\pi_+\mathfrak{D}_G\mathfrak{D}_F\pi_-R &= \pi_+\mathfrak{D}_G(\pi_+ + \pi_-)\mathfrak{D}_F\pi_-R \\ &= \pi_+\mathfrak{D}_G\pi_+\mathfrak{D}_F\pi_-R + \pi_+\mathfrak{D}_G\pi_-\mathfrak{D}_F\pi_-R \\ &= (\pi_+\mathfrak{D}_G\pi_+)(\pi_+\mathfrak{D}_F\pi_-R) + (\pi_+\mathfrak{D}_G\pi_-R)(R\pi_-\mathfrak{D}_F\pi_-R),\end{aligned} \tag{8.3}$$

where we have used that $R^2 = I$. The operators $\pi_+\mathfrak{D}_F\pi_-R$ and $\pi_+\mathfrak{D}_G\pi_-R$ are the Hankel operators of F and G respectively, and the operators $\pi_+\mathfrak{D}_G\pi_+$ and $R\pi_-\mathfrak{D}_F\pi_-R$ are bounded. If we choose F such that its Hankel operator belongs to S_p, then from (8.3) we conclude that the Hankel operator of GF belongs to S_p, as the Schatten classes are ideals in the space of bounded operators (see e.g. Böttcher & Silbermann [11, p.

14]). As such the bounded real singular values of GF belong to S_p for every $p > 0$ (and so are summable).

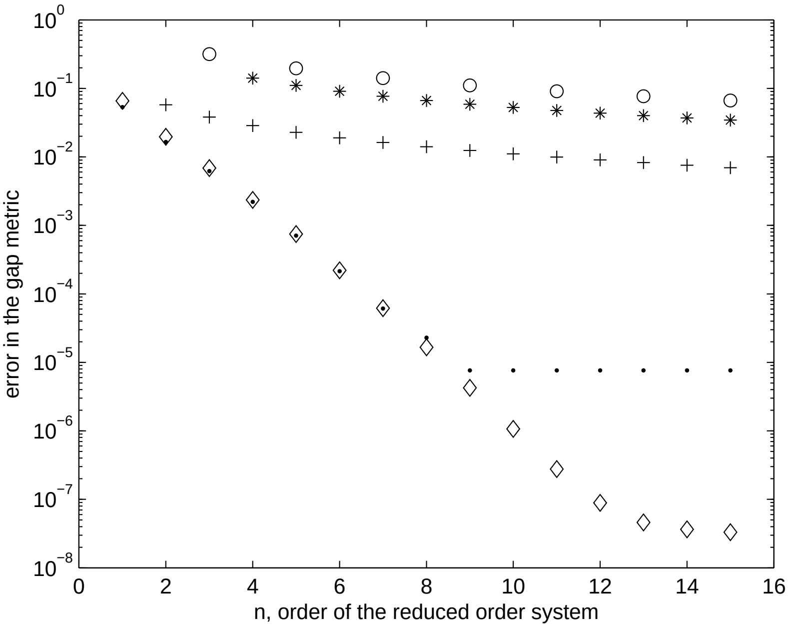
There are a variety of ways we can ensure that the Hankel operator of F belongs to S_p. For example, by Theorem 5.1.13, F is rational if and only if its Hankel operator is finite-rank, and in this instance belongs to S_p for every $p > 0$.

Remark 8.2.1. We remark that it seems to us that there are many infinite-dimensional, physically motivated systems that are bounded real (positive real), but that most are *not* strictly bounded real (positive real). We have tried to demonstrate, using the above example, that it is usually possible to tweak a bounded real (positive real) system to obtain a strictly bounded real (positive real) system.

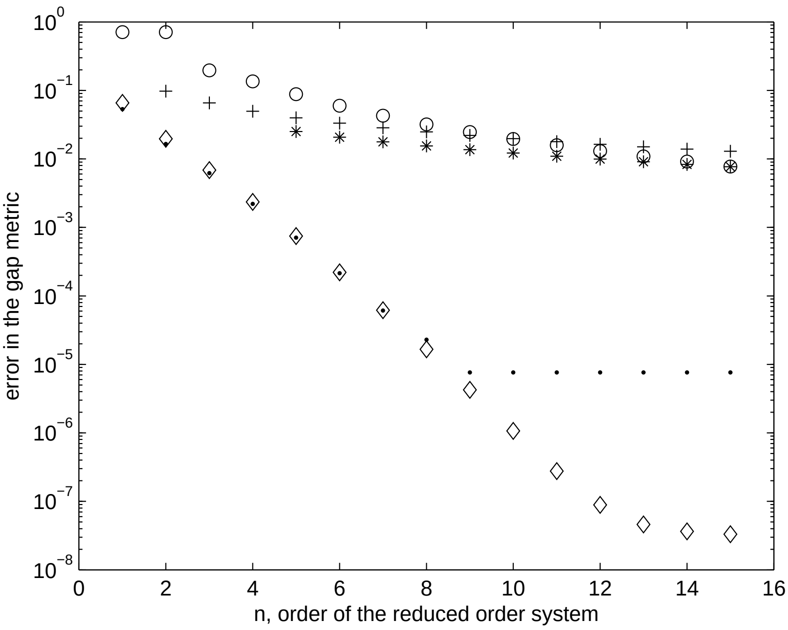
We have approximated the heat equation $(5.3)'$ (with $D = 1$) using several standard numerical approximation schemes. Unfortunately, computing the distance in the gap metric between these discretisations and the infinite-dimensional system is intractable. Consequently we have replaced the infinite-dimensional system with a linear finite-element approximation with $N = 50$ degrees of freedom, and have computed the corresponding gap metric distances using the MATLAB function `gapmetric`. The log of the gap metric error versus the order of the numerical discretisation is plotted in Figure 8-2.

Computing the positive real balanced truncation of the infinite-dimensional system is also intractable. Therefore we again take the linear finite-element approximation with $N = 50$ degrees of freedom as a substitute for the infinite-dimensional system and compute the positive real balanced truncation of this system. We note that this is the usual procedure for approximating balanced truncations of PDEs. Again, Figure 8-2 contains the log of the gap metric error between the positive real balanced truncation and the linear finite-element approximation with $N = 50$ versus the number of degrees of freedom in the positive real balanced truncation.

Figure 8-2 also contains the gap metric error bound for the positive real balanced truncation based on the positive real singular values of the linear finite-element approximation with $N = 50$ degrees of freedom. We see that the positive real balanced truncation converges much faster than the other schemes and that the error bound is very tight. Moreover, for $n \geq 8$ the error in the gap metric is larger than the error bound. This is consistent with the MATLAB function `gapmetric`, however, which is only accurate to 10^{-5}, which is the size of the errors. This suggests that for this example our gap metric error bound for $n \geq 8$ is in fact a better approximation of the actual error than the error computed by the `gapmetric` function in MATLAB.



(a)



(b)

Figure 8-2: *Approximation of heat equation* $(5.3)'$. *Both figures contain the positive real balanced truncation* $(\cdot)$ *and the gap metric error bound* $(\diamond)$. *Figure 8-2(a) in addition contains finite element approximations using linear* $(+)$, *quadratic* $(*)$ *and cubic* $(\circ)$ *elements. Figure 8-2(b) in addition contains finite difference approximations of order two* $(+)$ *and four* $(*)$ *and the Chebyshev collocation method* $(\circ)$.

8.3 Notes

The contents of this short chapter have been submitted (together with material from the previous two chapters) as [37]. The heat equation example and the core analysis from that example is borrowed from [61].

Chapter 9

Conclusions and future work

As mentioned in the introduction, the aim of this thesis was to extend bounded real and positive real balanced truncation in two different directions. In Part I we considered a more general concept of a dissipative system. We worked in the framework of dissipative driving-variable systems, and initially didn't differentiate between inputs and outputs. We did so for two reasons. Firstly so as to derive model reduction by balanced truncation in a framework free from the constrictions of input-output relations. We remark that although balanced truncation has been studied for behavioral systems (see the Notes section of Chapter 3), the gap metric error bound of Theorem 3.6.8 is new. Secondly, we sought to see bounded real and positive real balanced truncation of input-state-output systems as special cases of dissipative balanced truncation. As a corollary to the above gap metric error bound, we obtained the corresponding error bound for positive real balanced truncation, which is also new (as well as the less useful H^∞ error bound). The gap metric bound has been independently established by Timo Reis.

An obvious extension would be to consider infinite-dimensional versions of those results. The work on balanced truncation of improper positive real rational functions, started in Section 3.6.3, could also be developed further. Sufficient conditions for when the positive real balanced truncation exists in this instance are desirable.

In Part II we considered infinite-dimensional bounded real and positive real input-state-output systems. The picture here is nearly complete, in so much as we nearly have an analogous theory to that of Chapter 2. The main goals, we believe, were firstly to obtain rational approximants which preserved bounded realness or positive realness respectively. Secondly, to establish error bounds analogous to those in the finite-dimensional case, namely (2.18) and (2.32) for bounded real and positive real balanced truncation respectively, which we have achieved. We have also demonstrated that for some classes of infinite-dimensional systems, the positive real and bounded real balanced truncations converge mush faster than existing numerical schemes, justifying their use in model reduction of infinite-dimensional systems. The main unresolved issue is that we are forced to consider strictly bounded real and strictly positive real systems. It is sufficient in the finite-dimensional case that G (J) is bounded (positive) real for the error bound (2.18) ((2.32)) to hold; there we do not need to assume that G (J) is *strictly* bounded (positive) real. The example of Chapter 8 indicates to us that many bounded real and positive real systems that occur in physical situations are *not* strictly bounded real or strictly positive real respectively.

We have used the strictness assumption in construction of our extended systems, i.e. so that we can draw on the material of [97]. In the non-strict some results have been established by Curtain [16], and to what extent the results of [16] can be used as a substitute for [97] is to be investigated.

We would also like to build up some more examples of bounded real or positive real infinite-dimensional systems which our theory can be applied to, as lack of time has prohibited this so far. We have seen that parabolic PDEs with not too unbounded control and observation are often suitable.

Another important extension would be to address the issue of how to actually compute the bounded real and positive real balanced truncations. At the moment our results do not give a constructive method for this as the spectral factors cannot be found in general. It seems to us presently that for our theory to be applicable, some numerical approximation and convergence results are required. Theorem 5.0.3 is a result in this direction for Lyapunov balanced truncation and the discussion in Section 5.5.2 also considers this issue. The methods of [74] (and the references therein) seem a good starting point for further investigation.

Appendix

A

Proof of Lemma 3.6.13: The bottom route through the diagram gives

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{derived}} \begin{bmatrix} A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}} & BD_{\mathscr{U}}^{-1} \\ C_{\mathscr{Y}} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U}} & D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} \end{bmatrix} = \begin{bmatrix} A_{\mathrm{D}} & B_{\mathrm{D}} \\ C_{\mathrm{D}} & D_{\mathrm{D}} \end{bmatrix}.$$

Taking the spa we have

$$\begin{aligned}
(A_{\mathrm{D}})_{\mathrm{s}} &= (A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{11} - (A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{12}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{22}^{-1}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{21}, \\
(B_{\mathrm{D}})_{\mathrm{s}} &= B_1D_{\mathscr{U}}^{-1} - (A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{12}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{22}^{-1}B_2D_{\mathscr{U}}^{-1}, \\
(C_{\mathrm{D}})_{\mathrm{s}} &= (C_{\mathscr{Y},1} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},1}) \\
&\quad - (C_{\mathscr{Y},2} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},2})(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{22}^{-1}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{21}, \\
(D_{\mathrm{D}})_{\mathrm{s}} &= D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} - (C_{\mathscr{Y},2} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},2})(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{22}^{-1}B_2D_{\mathscr{U}}^{-1}.
\end{aligned}$$

The top route through the diagram gives

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \xrightarrow{\text{spa}} \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & B_1 - A_{12}A_{22}^{-1}B_2 \\ C_1 - C_2A_{22}^{-1}A_{21} & D - C_2A_{22}^{-1}B_2 \end{bmatrix} = \begin{bmatrix} A_{\mathrm{s}} & B_{\mathrm{s}} \\ C_{\mathrm{s}} & D_{\mathrm{s}} \end{bmatrix}.$$

Taking the derived $(\mathscr{U}, \mathscr{Y})$ system we have

$$\begin{aligned}
(A_{\mathrm{s}})_{\mathrm{D}} &= (A_{11} - A_{12}A_{22}^{-1}A_{21}) \\
&\quad - (B_1 - A_{12}A_{22}^{-1}B_2)(D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2)^{-1}(C_{\mathscr{U},1} - C_{\mathscr{U},2}A_{22}^{-1}A_{21}), \\
(B_{\mathrm{s}})_{\mathrm{D}} &= (B_1 - A_{12}A_{22}^{-1}B_2)(D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2)^{-1}, \\
(C_{\mathrm{s}})_{\mathrm{D}} &= (C_{\mathscr{Y},1} - C_{\mathscr{Y},2}A_{22}^{-1}A_{21}) \\
&\quad - (D_{\mathscr{Y}} - C_{\mathscr{Y},2}A_{22}^{-1}B_2)(D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2)^{-1}(C_{\mathscr{U},1} - C_{\mathscr{U},2}A_{22}^{-1}A_{21}), \\
(D_{\mathrm{s}})_{\mathrm{D}} &= (D_{\mathscr{Y}} - C_{\mathscr{Y},2}A_{22}^{-1}B_2)(D_{\mathscr{U}} - C_{\mathscr{U},2}A_{22}^{-1}B_2)^{-1}.
\end{aligned}$$

Note that both of these procedures are well-defined by our invertibility assumptions. It remains to show that

$$(A_{\mathrm{s}})_{\mathrm{D}} = (A_{\mathrm{D}})_{\mathrm{s}}, \quad (B_{\mathrm{s}})_{\mathrm{D}} = (B_{\mathrm{D}})_{\mathrm{s}}, \quad (C_{\mathrm{s}})_{\mathrm{D}} = (C_{\mathrm{D}})_{\mathrm{s}}, \quad (D_{\mathrm{s}})_{\mathrm{D}} = (D_{\mathrm{D}})_{\mathrm{s}}.$$

Set

$$L := A_{22} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2} \quad \text{and} \quad K := D_{\mathscr{U}} - C_{\mathscr{U},2} A_{22}^{-1} B_2,$$

which by assumption are both invertible. It is easy to see that

$$(A_{22} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2}) A_{22}^{-1} B_2 = B_2 D_{\mathscr{U}}^{-1} (D_{\mathscr{U}} - C_{\mathscr{U},2} A_{22}^{-1} B_2),$$

and hence

$$A_{22}^{-1} B_2 K^{-1} = L^{-1} B_2 D_{\mathscr{U}}^{-1}. \tag{9.1}$$

We will also need that

$$\begin{aligned} K^{-1} &= D_{\mathscr{U}}^{-1} + D_{\mathscr{U}}^{-1} C_{\mathscr{U},2} (A_{22} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2})^{-1} B_2 D_{\mathscr{U}}^{-1} \\ &= D_{\mathscr{U}}^{-1} + D_{\mathscr{U}}^{-1} C_{\mathscr{U},2} L^{-1} B_2 D_{\mathscr{U}}^{-1}. \end{aligned} \tag{9.2}$$

It now remains to calculate

$$\begin{aligned} (A_{\mathrm{s}})_{\mathrm{D}} &= (A_{11} - A_{12} A_{22}^{-1} A_{21}) - (B_1 - A_{12} A_{22}^{-1} B_2) K^{-1} (C_{\mathscr{U},1} - C_{\mathscr{U},2} A_{22}^{-1} A_{21}), \\ &= A_{11} - A_{12} A_{22}^{-1} A_{21} - B_1 K^{-1} C_{\mathscr{U},1} + A_{12} A_{22}^{-1} B_2 K^{-1} C_{\mathscr{U},1} \\ &\quad + B_1 K^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21} - A_{12} A_{22}^{-1} B_2 K^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21}. \end{aligned}$$

Using (9.1) in the fourth and sixth terms gives

$$\begin{aligned} (A_{\mathrm{s}})_{\mathrm{D}} &= A_{11} - A_{12} A_{22}^{-1} A_{21} - B_1 K^{-1} C_{\mathscr{U},1} + A_{12} L^{-1} B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},1} \\ &\quad + B_1 K^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21} - A_{12} L^{-1} B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21} \\ &= A_{11} - A_{12} L^{-1} (A_{21} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},1}) - B_1 K^{-1} C_{\mathscr{U},1} + B_1 K^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21} \\ &= (A - B D_{\mathscr{U}}^{-1} C_{\mathscr{U}})_{11} - (A - B D_{\mathscr{U}}^{-1} C_{\mathscr{U}})_{12} L^{-1} (A - B D_{\mathscr{U}}^{-1} C_{\mathscr{U}})_{21}, \qquad (9.3) \\ &= (A_{\mathrm{D}})_{\mathrm{s}}, \end{aligned}$$

where in (9.3) we have used (9.2) to compute that

$$\begin{aligned} -B_1 K^{-1} C_{\mathscr{U},1} + B_1 K^{-1} C_{\mathscr{U},2} A_{22}^{-1} A_{21} &= -B_1 D_{\mathscr{U}}^{-1} C_{\mathscr{U},1} \\ &\quad + B_1 D_{\mathscr{U}}^{-1} C_{\mathscr{U},2} L^{-1} (A_{21} - B_2 D_{\mathscr{U}}^{-1} C_{\mathscr{U},1}). \end{aligned}$$

The other calculations are similar. Using (9.2) we have

$$\begin{aligned}
(B_{\mathrm{s}})_{\mathrm{D}} &= (B_1 - A_{12}A_{22}^{-1}B_2)K^{-1} = (B_1 - A_{12}A_{22}^{-1}B_2)(D_{\mathscr{U}}^{-1} + D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}L^{-1}B_2D_{\mathscr{U}}^{-1}) \\
&= B_1D_{\mathscr{U}}^{-1} + B_1D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}L^{-1}B_2D_{\mathscr{U}}^{-1} - A_{12}A_{22}^{-1}[I + B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}L^{-1}]B_2D_{\mathscr{U}}^{-1} \\
&= B_1D_{\mathscr{U}}^{-1} + B_1D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}L^{-1}B_2D_{\mathscr{U}}^{-1} - A_{12}L^{-1}B_2D_{\mathscr{U}}^{-1} \\
&= B_1D_{\mathscr{U}}^{-1} - (A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{12}L^{-1}B_2D_{\mathscr{U}}^{-1} \\
&= (B_{\mathrm{D}})_{\mathrm{s}}.
\end{aligned}$$

$$\begin{aligned}
(C_{\mathrm{s}})_{\mathrm{D}} &= (C_{\mathscr{Y},1} - C_{\mathscr{Y},2}A_{22}^{-1}A_{21}) - (D_{\mathscr{Y}} - C_{\mathscr{Y},2}A_{22}^{-1}B_2)K^{-1}(C_{\mathscr{U},1} - C_{\mathscr{U},2}A_{22}^{-1}A_{21}) \\
&= C_{\mathscr{Y},1} - C_{\mathscr{Y},2}A_{22}^{-1}A_{21} - D_{\mathscr{Y}}K^{-1}C_{\mathscr{U},1} + C_{\mathscr{Y},2}A_{22}^{-1}B_2K^{-1}C_{\mathscr{U},1} \\
&\quad + D_{\mathscr{Y}}K^{-1}C_{\mathscr{U},2}A_{22}^{-1}A_{21} - C_{\mathscr{Y},2}A_{22}^{-1}B_2K^{-1}C_{\mathscr{U},2}A_{22}^{-1}A_{21}.
\end{aligned}$$

Using (9.1) in the fourth and sixth terms gives

$$\begin{aligned}
(C_{\mathrm{s}})_{\mathrm{D}} &= C_{\mathscr{Y},1} - C_{\mathscr{Y},2}A_{22}^{-1}A_{21} - D_{\mathscr{Y}}K^{-1}C_{\mathscr{U},1} + C_{\mathscr{Y},2}L^{-1}B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},1} \\
&\quad + D_{\mathscr{Y}}K^{-1}C_{\mathscr{U},2}A_{22}^{-1}A_{21} - C_{\mathscr{Y},2}L^{-1}B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}A_{22}^{-1}A_{21} \\
&= C_{\mathscr{Y},1} - C_{\mathscr{Y},2}L^{-1}(A_{21} - B_2D_{\mathscr{U}}^{-1}C_{\mathscr{U},1}) - D_{\mathscr{Y}}K^{-1}[C_{\mathscr{U},1} - C_{\mathscr{U},2}A_{22}^{-1}A_{21}] \\
&= C_{\mathscr{Y},1} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},1} - (C_{\mathscr{Y},2} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},2})L^{-1}(A - BD_{\mathscr{U}}^{-1}C_{\mathscr{U}})_{21}, \\
&= (C_{\mathrm{D}})_{\mathrm{s}}.
\end{aligned}$$

Finally, using (9.1) and (9.2) again we have

$$\begin{aligned}
(D_{\mathrm{s}})_{\mathrm{D}} &= (D_{\mathscr{Y}} - C_{\mathscr{Y},2}A_{22}^{-1}B_2)K^{-1} = D_{\mathscr{Y}}K^{-1} - C_{\mathscr{Y},2}A_{22}^{-1}B_2K^{-1} \\
&= D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} + D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},2}L^{-1}B_2D_{\mathscr{U}}^{-1} - C_{\mathscr{Y},2}L^{-1}B_2D_{\mathscr{U}}^{-1} \\
&= D_{\mathscr{Y}}D_{\mathscr{U}}^{-1} - (C_{\mathscr{Y},2} - D_{\mathscr{Y}}D_{\mathscr{U}}^{-1}C_{\mathscr{U},2})L^{-1}B_2D_{\mathscr{U}}^{-1} \\
&= (D_{\mathrm{D}})_{\mathrm{s}},
\end{aligned}$$

which completes the proof. □

B

Proof of Lemma 5.3.4: Since $(w_{i,k})_{i\in\mathbb{N}}^{1\le k\le p_i}$ is an orthonormal basis of $L^2(\mathbb{R}^+;\mathscr{Y})$, for any $f \in L^2(\mathbb{R}^+;\mathscr{Y})$ we have the decomposition

$$f = \sum_{i\in\mathbb{N}}\sum_{k=1}^{p_i}\langle w_{i,k}, f\rangle_{L^2}w_{i,k}.$$

For $n \in \mathbb{N}$ we obtain the orthogonal decomposition

$$L^2(\mathbb{R}^+;\mathscr{Y}) = \mathscr{X}_n^2 \oplus \mathscr{Z}_n^2, \tag{9.4}$$

where $\mathscr{X}_n^2$ is as stated in the lemma and $\mathscr{Z}_n^2 = (\mathscr{X}_n^2)^\perp$. We obtain orthogonal projections

$$\mathbb{P}_n : L^2(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n^2, \quad \mathbb{Q}_n := I - \mathbb{P}_n : L^2(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{Z}_n^2,$$

in the usual way so that

$$\mathbb{P}_n f := \sum_{i=1}^{n} \sum_{k=1}^{p_i} \langle w_{i,k}, f\rangle_{L^2} w_{i,k}. \tag{9.5}$$

Since $\mathscr{X}_n \subseteq W^{1,1} \subseteq L^2$ it follows that $\mathbb{P}_n$ restricts to a projection

$$P_n : W^{1,1}(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n,$$

which by (9.5), the triangle inequality and the inequalities

$$x \in W^{1,1}, \quad |\langle w_{i,k}, x\rangle_{L^2}| \le \|w_{i,k}\|_\infty \cdot \|x\|_1 \le \|w_{i,k}\|_{1,1} \cdot \|x\|_{1,1} < \infty,$$

is a bounded operator on $W^{1,1}(\mathbb{R}^+;\mathscr{Y})$ given by (5.73). Moreover, the range of P_n is all of $\mathscr{X}_n$. Therefore $\mathbb{Q}_n$ restricts to a bounded projection

$$Q_n := I - P_n, \quad \text{on } W^{1,1}(\mathbb{R}^+;\mathscr{Y}).$$

Define $\mathscr{Z}_n^{1,1} := Q_n W^{1,1}(\mathbb{R}^+;\mathscr{Y}) \subseteq W^{1,1}(\mathbb{R}^+;\mathscr{Y})$ so that

$$Q_n : W^{1,1}(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{Z}_n^{1,1}, \quad \mathscr{Z}_n^{1,1} \subseteq \mathscr{Z}_n^2$$

and

$$W^{1,1}(\mathbb{R}^+;\mathscr{Y}) = \mathscr{X}_n^{1,1} \oplus \mathscr{Z}_n^{1,1}.$$

We see that $\mathscr{X}_n^{1,1}$ and $\mathscr{Z}_n^{1,1}$ are orthogonal in the L^2 sense because $\mathscr{X}_n^{1,1}$ and $\mathscr{Z}_n^2$ are. The projection P_n defined by (5.73) on $W^{1,1}$ extends to a bounded projection

$$\mathcal{P}_n : L^1(\mathbb{R}^+;\mathscr{Y}) \to \mathscr{X}_n^2,$$

given by (5.73) where again the L^2 inner product is the duality product between L^∞ and L^1, because

$$x \in L^1, \quad |\langle w_{i,k}, x\rangle_{L^2}| \le \|w_{i,k}\|_\infty \cdot \|x\|_1 \le \|w_{i,k}\|_{1,1} \cdot \|x\|_1 < \infty.$$

Since $W^{1,1}$ is dense in L^1 it follows that $\mathcal{P}_n$ defined by (5.73) must be the unique continuous extension of P_n to L^1. We can now extend Q_n to a bounded projection

$$\mathcal{Q}_n := I - \mathcal{P}_n, \quad \text{on } L^1(\mathbb{R}^+; \mathscr{Y}).$$

Define $\mathscr{Z}_n^1 := \mathcal{Q}_n L^1(\mathbb{R}^+; \mathscr{Y})$ so that

$$\mathcal{Q}_n : L^1(\mathbb{R}^+; \mathscr{Y}) \to \mathscr{Z}_n^1, \quad \mathscr{Z}_n^{1,1} \subseteq \mathscr{Z}_n^1,$$

and

$$L^1(\mathbb{R}^+; \mathscr{Y}) = \mathscr{X}_n^1 \oplus \mathscr{Z}_n^1.$$

We claim that $\mathscr{X}_n^1$ and $\mathscr{Z}_n^1$ are orthogonal with respect to the duality-product. So we seek to prove

$$\langle x, y \rangle_{L^2} = 0, \quad \forall\, x \in \mathscr{X}_n^1, \quad \forall\, y \in \mathscr{Z}_n^1. \tag{9.6}$$

Let $x \in \mathscr{X}_n^2$ and let $y \in \mathscr{Z}_n^1$. Then y is the L^1 limit of a sequence $(y_k)_{k\in\mathbb{N}} \subseteq W^{1,1}$. Note that as $\mathcal{Q}_n$ is bounded

$$\mathcal{Q}_n y_k \xrightarrow{L^1} \mathcal{Q}_n y = y,$$

and as $\mathcal{Q}_n$ extends Q_n

$$\mathcal{Q}_n y_k = Q_n y_k \in \mathscr{Z}_n^{1,1}.$$

Therefore y is the L^1 limit of a sequence in $\mathscr{Z}_n^{1,1}$. We have already seen that

$$\langle x, y \rangle_{L^2} = 0, \quad \forall\, x \in \mathscr{X}_n, \quad \forall\, y \in \mathscr{Z}_n^{1,1},$$

which implies that

$$\begin{aligned} |\langle x, y \rangle_{L^2}| = |\langle x, y \rangle_{L^2} - \underbrace{\langle x, Q_n y_k \rangle_{L^2}}_{=0}| &= |\langle x, y - Q_n y_k \rangle_{L^2}| \le \|x\|_\infty \cdot \|y - Q_n y_k\|_1 \\ &\le \|x\|_{1,1} \cdot \|y - Q_n y_k\|_1 \to 0, \quad \text{as} \quad k \to \infty. \end{aligned}$$

We conclude that (9.6) holds. It remains to prove the self-adjoint properties in (5.74) and (5.75). Both sides of the first equation in (5.74) are well-defined as

$$\mathcal{P}_n x, \mathcal{P}_n y \in \mathscr{X}_n^1 \subseteq W^{1,1} \subseteq L^\infty.$$

We have for $x, y \in L^1$

$$\begin{aligned} \langle x, \mathcal{P}_n y \rangle_{L^2} &= \langle \mathcal{P}_n x, \mathcal{P}_n y \rangle_{L^2} + \langle (I - \mathcal{P}_n) x, \mathcal{P}_n y \rangle_{L^2} \\ &= \langle \mathcal{P}_n x, \mathcal{P}_n y \rangle_{L^2} + \langle \mathcal{Q}_n x, \mathcal{P}_n y \rangle_{L^2}. \end{aligned}$$

As $\mathcal{Q}_n x \in \mathscr{Z}_n^1$ and $\mathcal{P}_n y \in \mathscr{X}_n^1$, the orthogonality of $\mathscr{X}_n^1$ and $\mathscr{Z}_n^1$ established above implies that $\langle \mathcal{Q}_n x, \mathcal{P}_n y\rangle_{L^2} = 0$. Thus

$$\langle x, \mathcal{P}_n y\rangle_{L^2} = \langle \mathcal{P}_n x, \mathcal{P}_n y\rangle_{L^2},$$

which is symmetric in x and y. We conclude that (5.74) holds. The proof of the second equation in (5.74) (involving $\mathcal{Q}_n$) is identical to that above with $\mathcal{P}_n$ and $\mathcal{Q}_n$ interchanged. To see (5.75) is true, we argue similarly. For $x, y \in W^{1,1}$

$$\begin{aligned}\langle x, P_n y\rangle_{L^2} &= \langle P_n x, P_n y\rangle_{L^2} + \langle (I - P_n)x, P_n y\rangle_{L^2}\\ &= \langle P_n x, P_n y\rangle_{L^2} + \langle Q_n x, P_n y\rangle_{L^2}.\end{aligned}$$

As $Q_n x \in \mathscr{Z}_n^{1,1}$ and $P_n y \in \mathscr{X}_n^{1,1}$, the orthogonality of $\mathscr{X}_n^{1,1}$ and $\mathscr{Z}_n^{1,1}$ established above implies that $\langle Q_n x, P_n y\rangle_{L^2} = 0$. Thus

$$\langle x, P_n y\rangle_{L^2} = \langle P_n x, P_n y\rangle_{L^2},$$

which again is symmetric in x and y and so we infer (5.75) as before. The proof for Q_n follows similarly (by interchanging Q_n and P_n). $\square$

Bibliography

[1] Vadym M. Adamjan, Damir Z. Arov, and Mark G. Kreĭn. Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem. *Mat. Sb. (N.S.)*, 86(128):34–75, 1971.

[2] Brian D.O. Anderson and Sumeth Vongpanitlerd. *Network analysis and synthesis: a modern systems theory approach.* Prentice Hall, 1973.

[3] Athanasios C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[4] Damir Z. Arov and Olof J. Staffans. State/signal linear time-invariant systems theory. I. Discrete time systems. In *The state space method generalizations and applications*, volume 161 of *Oper. Theory Adv. Appl.*, pages 115–177. Birkhäuser, Basel, 2006.

[5] Damir Z. Arov and Olof J. Staffans. The infinite-dimensional continuous time Kalman-Yakubovich-Popov inequality. In *The extended field of operator theory*, volume 171 of *Oper. Theory Adv. Appl.*, pages 37–72. Birkhäuser, Basel, 2007.

[6] Damir Z. Arov and Olof J. Staffans. State/signal linear time-invariant systems theory. II. Passive discrete time systems. *Internat. J. Robust Nonlinear Control*, 17(5-6):497–548, 2007.

[7] Damir Z. Arov and Olof J. Staffans. State/signal linear time-invariant systems theory. III. Transmission and impedance representations of discrete time systems. In *Operator theory, structured matrices, and dilations*, volume 7 of *Theta Ser. Adv. Math.*, pages 29–68. Theta, Bucharest, 2007.

[8] Damir Z. Arov and Olof J. Staffans. State/signal linear time-invariant systems theory. IV. Affine and generalized transfer functions of discrete time systems. *Complex Anal. Oper. Theory*, 1(4):457–521, 2007.

[9] Vitold Belevitch. *Classical network theory.* Holden-Day, San Francisco, Calif., 1968.

[10] János Bognár. *Indefinite inner product spaces.* Springer-Verlag, New York, 1974. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78.

[11] Albrecht Böttcher and Bernd Silbermann. *Analysis of Toeplitz operators.* Springer-Verlag, Berlin, 1990.

[12] Françoise Chatelin. *Spectral approximation of linear operators.* Computer Science and Applied Mathematics. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983.

[13] Xiang Chen and Kemin Zhou. On the relative and multiplicative model reductions. In *Proceedings of the 27th Southeastern Symposium on System Theory (SSST'95)*, page 57. IEEE Computer Society, 1995.

[14] Xin Chen and John T. Wen. Positive realness preserving model reduction with $\mathcal{H}^\infty$ norm error bounds. *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, 42(1):23–29, 1995.

[15] Ronald R. Coifman and Richard Rochberg. Representation theorems for holomorphic and harmonic functions in L^p. In *Representation theorems for Hardy spaces*, volume 77 of *Astérisque*, pages 11–66. Soc. Math. France, Paris, 1980.

[16] Ruth F. Curtain. Linear operator inequalities for strongly stable weakly regular linear systems. *Math. Control Signals Systems*, 14(4):299–337, 2001.

[17] Ruth F. Curtain and Keith Glover. Balanced realisations for infinite-dimensional systems. In *Operator theory and systems (Amsterdam, 1985)*, volume 19 of *Oper. Theory Adv. Appl.*, pages 87–104. Birkhäuser, Basel, 1986.

[18] Ruth F. Curtain, Kalle Mikkola, and Amol J. Sasane. The Hilbert-Schmidt property of feedback operators. *J. Math. Anal. Appl.*, 329(2):1145–1160, 2007.

[19] Ruth F. Curtain and Amol J. Sasane. Compactness and nuclearity of the Hankel operator and internal stability of infinite-dimensional state linear systems. *Internat. J. Control*, 74(12):1260–1270, 2001.

[20] Ruth F. Curtain and Hans Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995.

[21] Liyi Dai. *Singular control systems*, volume 118 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1989.

[22] Uday B. Desai and Debajyoti Pal. A transformation approach to stochastic model reduction. *IEEE Trans. Automat. Control*, 29(12):1097–1100, 1984.

[23] Dale F. Enns. Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *Proc. CDC*, pages 127–132, 1984. doi: 10.1109/CDC.1984.272286.

[24] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[25] John B. Garnett. *Bounded analytic functions*, volume 236 of *Graduate Texts in Mathematics*. Springer, New York, first edition, 2007.

[26] Keith Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Internat. J. Control*, 39(6):1115–1193, 1984.

[27] Keith Glover, Ruth F. Curtain, and Jonathan R. Partington. Realisation and approximation of linear infinite-dimensional systems with error bounds. *SIAM J. Control Optim.*, 26(4):863–898, 1988.

[28] Keith Glover, James Lam, and Jonathan R. Partington. Rational approximation of a class of infinite-dimensional systems. I. Singular values of Hankel operators. *Math. Control Signals Systems*, 3(4):325–344, 1990.

[29] Israel Gohberg, Seymour Goldberg, and Marinus A. Kaashoek. *Classes of linear operators. Vol. I*, volume 49 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1990.

[30] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrices and indefinite scalar products*, volume 8 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1983.

[31] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Indefinite linear algebra and applications*. Birkhäuser Verlag, Basel, 2005.

[32] Michael Green. A relative error bound for balanced stochastic truncation. *IEEE Trans. Automat. Control*, 33(10):961–965, 1988.

[33] Michael Green and David J. N. Limebeer. *Linear robust control*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.

[34] Serkan Gugercin and Athanasios C. Antoulas. A survey of model reduction by balanced truncation and some new results. *Internat. J. Control*, 77(8):748–766, 2004.

[35] Chris Guiver and Mark R. Opmeer. A counter-example to "positive realness preserving model reduction with $\mathcal{H}_\infty$ norm error bounds". *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 58(6):1410–1411, June 2011.

[36] Chris Guiver and Mark R. Opmeer. Model reduction by balanced truncation for systems with nuclear Hankel operators. Submitted, 2011.

[37] Chris Guiver and Mark R. Opmeer. Bounded real and positive real balanced truncation for infinite-dimensional systems. Submitted, 2012.

[38] Chris Guiver and Mark R. Opmeer. An error bound in the gap metric for dissipative balanced approximations. Submitted, 2012.

[39] P. Harshavardhana. *Model Reduction methods in control and signal processing.* PhD thesis, Univ. South. Calif. EE Dept, Los Angeles, CA, 1984.

[40] P. Harshavardhana, Edmond A. Jonckheere, and Leonard M. Silverman. Stochastic balancing and approximation-stability and minimality. *IEEE Trans. Automat. Control*, 29(8):744–746, 1984.

[41] Edmond A. Jonckheere and Leonard M. Silverman. A new set of invariants for linear systems—application to reduced order compensator design. *IEEE Trans. Automat. Control*, 28(10):953–964, 1983.

[42] Leonid V. Kantorovich and Gleb P. Akilov. *Functional analysis.* Pergamon Press, Oxford, second edition, 1982. Translated from the Russian by Howard L. Silcock.

[43] Tosio Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[44] Mikael Kurula and Olof J. Staffans. Well-posed state/signal systems in continuous time. *Complex Anal. Oper. Theory*, 4(2):319–390, 2010.

[45] Mikael Kurula and Olof J. Staffans. Connections between classical and generalised trajectories of a state/signal system. *Complex Anal. Oper. Theory*, 5(2):403–422, 2011.

[46] Peter D. Lax. Symmetrizable linear transformations. *Comm. Pure Appl. Math.*, 7:633–647, 1954.

[47] Peter D. Lax. *Functional analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002.

[48] Yi Liu and Brian D. O. Anderson. Controller reduction via stable factorization and balancing. *Internat. J. Control*, 44(2):507–531, 1986.

[49] David G. Meyer. Fractional balanced reduction: model reduction via fractional representation. *IEEE Trans. Automat. Control*, 35(12):1341–1345, 1990.

[50] Kalle Mikkola. *Infinite-dimensional Linear Systems, Optimal Control and Algebraic Riccati Equations.* Research reports / Helsinki University of Technology, Institute of Mathematics. Helsinki University of Technology, 2002.

[51] Ha Binh Minh. *Model Reduction in a Behavioral Framework.* PhD thesis, Groningen, 2009.

[52] Bruce C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, 26(1):17–32, 1981.

[53] Giovanni Muscato, G. Nunnari, and Luigi Fortuna. Singular perturbation approximation of bounded real balanced and stochastically balanced transfer matrices. *Internat. J. Control*, 66(2):253–269, 1997.

[54] Nikolai K. Nikolski. *Operators, functions, and systems: an easy reading. Vol. 1*, volume 92 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2002. Hardy, Hankel, and Toeplitz, Translated from the French by Andreas Hartmann.

[55] Raimund J. Ober. A note on a system theoretic approach to a conjecture by Peller-Khrushchev. *Systems Control Lett.*, 8(4):303–306, 1987.

[56] Raimund J. Ober and Stephen Montgomery-Smith. Bilinear transformation of infinite-dimensional state-space systems and balanced realizations of nonrational transfer functions. *SIAM J. Control Optim.*, 28(2):438–465, 1990.

[57] Philippe C. Opdenacker and Edmond A. Jonckheere. A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds. *IEEE Trans. Circuits and Systems*, 35(2):184–189, 1988.

[58] Mark R. Opmeer. Model reduction for distributed parameter systems: a functional analytic view. Accepted, American Control Conference, 2012.

[59] Mark R. Opmeer. Infinite-dimensional linear systems: a distributional approach. *Proc. London Math. Soc. (3)*, 91(3):738–760, 2005.

[60] Mark R. Opmeer. *Model reduction for controller design for infinite-dimensional systems.* PhD thesis, Groningen, 2006.

[61] Mark R. Opmeer. Decay of Hankel singular values of analytic control systems. *Systems Control Lett.*, 59(10):635–638, 2010.

[62] Lavon B. Page. Bounded and compact vectorial Hankel operators. *Trans. Amer. Math. Soc.*, 150:529–539, 1970.

[63] Jonathan R. Partington. *An introduction to Hankel operators*, volume 13 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1988.

[64] Jonathan R. Partington. *Linear operators and linear systems. An analytical approach to control theory*, volume 60 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2004.

[65] Vladimir V. Peller. *Hankel operators and their applications*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[66] Lars Pernebo and Leonard M. Silverman. Model reduction via balanced state space representations. *IEEE Trans. Automat. Control*, 27(2):382–387, 1982.

[67] Jan Willem Polderman and Jan C. Willems. *Introduction to mathematical systems theory*, volume 26 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1998. A behavioral approach.

[68] Richard Rebarber. Conditions for the equivalence of internal and external stability for distributed parameter systems. *IEEE Trans. Automat. Control*, 38(6):994–998, 1993.

[69] Timo Reis. Private correspondence, 2011.

[70] Timo Reis and Tatjana Stykel. Positive real and bounded real balancing for model reduction of descriptor systems. *International Journal of Control*, 83(1):74–88, 2010.

[71] Marvin Rosenblum and James Rovnyak. *Hardy classes and operator theory*. Dover Publications Inc., Mineola, NY, 1997. Corrected reprint of the 1985 original.

[72] Dietmar Salamon. Infinite-dimensional linear systems with unbounded control and observation: a functional analytic approach. *Trans. Amer. Math. Soc.*, 300(2):383–431, 1987.

[73] Dietmar Salamon. Realization theory in Hilbert space. *Math. Systems Theory*, 21(3):147–164, 1989.

[74] John R. Singler. Snapshot algorithm for balanced model reduction of linear parabolic systems: convergence theory. *Submitted for publication*, 2009.

[75] Y.L. Šmuljan. Invariant subspaces of semigroups and the lax–phillips scheme. *Deposited in VINITI*, (8009-B86), 1986.

[76] Olof J. Staffans. Quadratic optimal control of regular well-posed linear systems, with applications to parabolic equations, 1997.

[77] Olof J. Staffans. Quadratic optimal control of stable well-posed linear systems. *Trans. Amer. Math. Soc.*, 349(9):3679–3715, 1997.

[78] Olof J. Staffans. Quadratic optimal control of well-posed linear systems. *SIAM J. Control Optim.*, 37(1):131–164 (electronic), 1999.

[79] Olof J. Staffans. Passive and conservative continuous-time impedance and scattering systems. I. Well-posed systems. *Math. Control Signals Systems*, 15(4):291–315, 2002.

[80] Olof J. Staffans. Passive and conservative infinite-dimensional impedance and scattering systems (from a personal point of view). In *Mathematical systems theory in biology, communications, computation, and finance (Notre Dame, IN, 2002)*, volume 134 of *IMA Vol. Math. Appl.*, pages 375–413. Springer, New York, 2003.

[81] Olof J. Staffans. *Well-posed linear systems*, volume 103 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.

[82] Weiqian Sun, Pramod P. Khargonekar, and Duksun Shim. Solution to the positive real control problem for linear time-invariant systems. *IEEE Trans. Automat. Control*, 39(10):2034–2046, 1994.

[83] Charles Swartz. *An introduction to functional analysis*, volume 157 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1992.

[84] Sergei R. Treil′. An inverse spectral problem for the modulus of the Hankel operator, and balanced realizations. *Algebra i Analiz*, 2(2):158–182, 1990.

[85] Harry L. Trentelman. Bounded real and positive real balanced truncation using Sigma-normalised coprime factors. *Systems Control Lett.*, 58(12):871 – 879, 2009.

[86] Harry L. Trentelman. A behavioral characterization of the positive real and bounded real characteristic values in balancing. *Systems Control Lett.*, 60(1):7–14, 2011.

[87] Andras Varga. Computation of coprime factorizations of rational matrices. In *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, volume 5, pages 4830 –4835 vol.5, dec 1997.

[88] Andras Varga. Computation of normalized coprime factorizations of rational matrices. *Systems Control Lett.*, 33(1):37–45, 1998.

[89] Erik I. Verriest. Low sensitivity design and optimal order reduction for the LQG-problem. In *Proc. 24th Midwest Symp. Circuits Syst.*, pages 365–369, 1981.

[90] Weizheng Wang and Michael G. Safonov. A tighter relative-error bound for balanced stochastic truncation. *Systems Control Lett.*, 14(4):307–317, 1990.

[91] Siep Weiland. A behavioral approach to balanced representations of dynamical systems. *Linear Algebra Appl.*, 205/206:1227–1252, 1994.

[92] George Weiss. Admissibility of unbounded control operators. *SIAM J. Control Optim.*, 27(3):527–545, 1989.

[93] George Weiss. Admissible observation operators for linear semigroups. *Israel J. Math.*, 65(1):17–43, 1989.

[94] George Weiss. Representation of shift-invariant operators on L^2 by H^∞ transfer functions: an elementary proof, a generalization to L^p, and a counterexample for L^∞. *Math. Control Signals Systems*, 4(2):193–203, 1991.

[95] George Weiss. Transfer functions of regular linear systems. I. Characterizations of regularity. *Trans. Amer. Math. Soc.*, 342(2):827–854, 1994.

[96] George Weiss and Hans Zwart. An example in linear quadratic optimal control. *Systems Control Lett.*, 33(5):339–349, 1998.

[97] Martin Weiss and George Weiss. Optimal control of stable weakly regular linear systems. *Math. Control Signals Systems*, 10(4):287–330, 1997.

[98] John T. Wen. Time domain and frequency domain conditions for strict positive realness. *IEEE Trans. Automat. Control*, 33(10):988–992, 1988.

[99] Jan C. Willems. Dissipative dynamical systems. II. Linear systems with quadratic supply rates. *Arch. Rational Mech. Anal.*, 45:352–393, 1972.

[100] Jan C. Willems. Input-output and state-space representations of finite-dimensional linear time-invariant systems. *Linear Algebra Appl.*, 50:581–608, 1983.

[101] Jan C. Willems. The behavioral approach to open and interconnected systems: modeling by tearing, zooming, and linking. *IEEE Control Syst. Mag.*, 27(6):46–99, 2007.

[102] Jan C. Willems and Harry L. Trentelman. Synthesis of dissipative systems using quadratic differential forms. I. *IEEE Trans. Automat. Control*, 47(1):53–69, 2002.

[103] Nicholas J. Young. Balanced realizations in infinite dimensions. *Operator Theory: Advances and Applications*, 19:449–471, 1986.

[104] Kemin Zhou, John C. Doyle, and Keith Glover. *Robust and optimal control*. Prentice Hall Englewood Cliffs, NJ, 1996.